

On chemical distances and shape theorems in percolation models with long-range correlations

Alexander Drewitz *

Balázs Ráth †

Artëm Sapozhnikov ‡

December 13, 2012

Abstract

In this paper we provide general conditions on a one parameter family of random infinite subsets of \mathbb{Z}^d to contain a unique infinite connected component for which the chemical distances are comparable to the Euclidean distances, focusing primarily on models with long-range correlations. Our results are in the spirit of [2] proved for Bernoulli percolation. We also prove a shape theorem for balls in the chemical distance under such conditions. Our general statements give novel results about the structure of the infinite connected component of the vacant set of random interlacements and the level sets of the Gaussian free field. We also obtain alternative proofs to the main results in [8]. Finally, as a corollary, we obtain new results about the (chemical) diameter of the largest connected component in the complement of the trace of the random walk on the torus.

Percolation was introduced in [7] as a mathematical model of a porous medium, where the physical space is modeled by the lattice \mathbb{Z}^d , and the pure substance is described as a random subset \mathcal{S} of \mathbb{Z}^d . A central question is to understand physical properties of \mathcal{S} , and how universal these properties are for various distributions of \mathcal{S} . Mathematically speaking, what can be said about the connectivity properties and the geometry of the subgraph of \mathbb{Z}^d spanned by \mathcal{S} ? In this paper we consider the situation when the set \mathcal{S} has a unique infinite connected component \mathcal{S}_∞ which covers a positive fraction of \mathbb{Z}^d . We define the *chemical distance* of $x, y \in \mathcal{S}_\infty$ to be the length of the shortest nearest neighbor path connecting x and y in \mathcal{S}_∞ , and we want to understand when the long-scale behavior of the resulting random metric space is similar to that of the underlying space \mathbb{Z}^d .

In the case of supercritical Bernoulli percolation (when every vertex of \mathbb{Z}^d is in \mathcal{S} with probability $p > p_c$ independently of each other), [2, Theorem 1.1] asserts the existence of a constant $C = C(d, p)$ such that the probability that the chemical distance of $x, y \in \mathcal{S}_\infty$ is greater than $C|x - y|$ decays exponentially as $|x - y| \rightarrow \infty$. Using this control on chemical distances as well as Kingman's subadditive ergodic theorem one can deduce a *shape theorem* which states that the chemical ball of radius R , rescaled by R ,

*Columbia University, Department of Mathematics, RM 614, MC 4419, 2990 Broadway, New York City, NY 10027, USA. email: drewitz@math.columbia.edu

†The University of British Columbia, Department of Mathematics, Room 121, 1984 Mathematics Road, Vancouver, B.C. Canada V6T 1Z2. email: rathb@math.ubc.ca

‡Max-Planck Institute for Mathematics in the Sciences, Inselstrasse 22, 04103 Leipzig, Germany. email: artem.sapozhnikov@mis.mpg.de

⁰MSC2000: Primary 60K35, 82B43.

⁰Keywords: Chemical Distance, Shape Theorem, Percolation, Long-Range Correlations, Random Interlacements, Gaussian Free Field.

almost surely converges with respect to the Hausdorff distance to a deterministic compact convex set of \mathbb{R}^d as $R \rightarrow \infty$ (see, e.g., [11, Corollary 5.4]).

In this paper we extend the result of [2] to a general class of random subsets \mathcal{S} , focusing primarily on models with long-range correlations. We provide sufficient conditions on the distribution of \mathcal{S} to satisfy analogues to results of [2]. These conditions are general enough to be satisfied by many models, including random interacements, the vacant set of random interacements, and the level sets of the Gaussian free field (see Sections 2.2, 2.3, and 2.4, respectively). For the last two models our general statements give novel results. Analogous questions about random interacements were answered in [8], and our method provides an alternative, model independent approach to the results of [8]. Finally, using a connection between random interacements and random walk on the discrete torus, we obtain new results about the (chemical) diameter of the largest connected component in the complement of the random walk trace (see Section 2.5).

1 Model and results

We consider a one parameter family of probability measures \mathbb{P}^u , $u \in (a, b) \subseteq \mathbb{R}_+$, on the measurable space $(\{0, 1\}^{\mathbb{Z}^d}, \mathcal{F})$, $d \geq 2$, where the sigma-algebra \mathcal{F} is generated by the canonical coordinate maps $\Psi_x : \{0, 1\}^{\mathbb{Z}^d} \rightarrow \{0, 1\}$, $x \in \mathbb{Z}^d$ (i.e., $\Psi_x(\xi) = \xi_x$ for $\xi \in \{0, 1\}^{\mathbb{Z}^d}$ and $x \in \mathbb{Z}^d$).

For any $\xi \in \{0, 1\}^{\mathbb{Z}^d}$, we define

$$\mathcal{S} = \mathcal{S}(\xi) = \{x \in \mathbb{Z}^d : \xi_x = 1\} \subseteq \mathbb{Z}^d.$$

We view \mathcal{S} as a subgraph of \mathbb{Z}^d in which the edges are drawn between any two vertices of \mathcal{S} within ℓ^1 -distance 1 from each other. (For $x = (x^{(1)}, \dots, x^{(d)}) \in \mathbb{R}^d$, the ℓ^1 -norm of x is defined in the usual way by $|x|_1 = \sum_{i=1}^d |x^{(i)}|$.) We denote by \mathcal{S}_∞ the subset of vertices of \mathcal{S} which are in infinite connected components of \mathcal{S} .

In this paper we will focus on connectivity properties of \mathcal{S}_∞ . We will prove that under certain conditions (see **P1** – **P3** and **S1** – **S2**) on the family of probability measures \mathbb{P}^u , $u \in (a, b)$, the graph \mathcal{S}_∞ is connected and “looks like” the underlying lattice \mathbb{Z}^d . Roughly speaking, we show that on large scales the graph distance in \mathcal{S}_∞ behaves like a metric induced by a norm on \mathbb{R}^d .

We will now define conditions on the family \mathbb{P}^u , $u \in (a, b)$. After that we will state the main results of the paper. Particular examples of probability measures satisfying the conditions (and results) below will be given in Section 2.

The numbers $0 \leq a < b$ as well as the dimension $d \geq 2$ are going to be fixed throughout the paper, and we omit the dependence of various constants on a , b , and d .

P1 For any $u \in (a, b)$, \mathbb{P}^u is invariant and ergodic with respect to the lattice shifts.

To state the next two conditions we need the following definition.

An event $G \in \mathcal{F}$ is called *increasing* (respectively, *decreasing*), if for all $\xi \in G$ and $\xi' \in \{0, 1\}^{\mathbb{Z}^d}$ with $\xi_y \leq \xi'_y$ (respectively, $\xi_y \geq \xi'_y$) for all $y \in \mathbb{Z}^d$, one has $\xi' \in G$.

P2 For any $u, u' \in (a, b)$ with $u < u'$, and any increasing event $G \in \mathcal{F}$, $\mathbb{P}^u[G] \leq \mathbb{P}^{u'}[G]$. (Usually, this condition is referred to as stochastic monotonicity of \mathbb{P}^u .)

Conditions **P1** and **P2** are rather general and are satisfied by many families of probability measures. The condition **P3** below is more restrictive than the above two. It states that \mathbb{P}^u , $u \in (a, b)$, satisfy a certain weak decorrelation inequality. For $x \in \mathbb{Z}^d$ and $r \in \mathbb{R}_+$, we denote by

$$B(x, r) = \{y \in \mathbb{Z}^d : |x - y|_\infty \leq \lfloor r \rfloor\}$$

the closed l^∞ -ball in \mathbb{Z}^d with radius $\lfloor r \rfloor$ and center x . (For $x = (x^{(1)}, \dots, x^{(d)}) \in \mathbb{R}^d$, the l^∞ -norm of x is defined by $|x| = |x|_\infty = \max\{|x^{(1)}|, \dots, |x^{(d)}|\}$.) We write $B(r)$ for $B(0, r)$.

P3 Let $L \geq 1$ be an integer. Let $x_1, x_2 \in \mathbb{Z}^d$. For $i \in \{1, 2\}$, let $A_i \in \sigma(\Psi_y : y \in B(x_i, 10L))$ be decreasing events, and $B_i \in \sigma(\Psi_y : y \in B(x_i, 10L))$ increasing events. There exist $R_P, L_P < \infty$ and $\varepsilon_P, \chi_P > 0$ such that for any integer $R \geq R_P$ and $a < \hat{u} < u < b$ satisfying

$$u \geq (1 + R^{-\chi_P}) \cdot \hat{u}, \quad (1.1)$$

if $|x_1 - x_2|_\infty \geq R \cdot L$, then

$$\mathbb{P}^u[A_1 \cap A_2] \leq \mathbb{P}^{\hat{u}}[A_1] \cdot \mathbb{P}^{\hat{u}}[A_2] + e^{-f_P(L)}, \quad (1.2)$$

and

$$\mathbb{P}^{\hat{u}}[B_1 \cap B_2] \leq \mathbb{P}^u[B_1] \cdot \mathbb{P}^u[B_2] + e^{-f_P(L)}, \quad (1.3)$$

where f_P is a real valued function satisfying

$$f_P(L) \geq e^{(\log L)^{\varepsilon_P}} \text{ for all } L \geq L_P. \quad (1.4)$$

Remark 1.1. At first sight it may look like condition **P3** is rather strong and asserts that the correlations should decay much faster than any polynomial ($e^{f_P(L)} \gg L^k$ for any k). However, the fact that we deal with different parameters u and \hat{u} (and a very restricted class of events) allows for measures with polynomial decay of correlations to satisfy **P3**. Examples of such families of measures for $d \geq 3$ are random interacements and the level sets of the Gaussian free field, for which

$$|\mathbb{P}^u[\Psi_x = 1, \Psi_y = 1] - \mathbb{P}^u[\Psi_x = 1]\mathbb{P}^u[\Psi_y = 1]| \asymp (1 + |x - y|)^{2-d}, \quad x, y \in \mathbb{Z}^d.$$

We will discuss these examples in more details in Sections 2.2 – 2.4. In the literature, the idea behind the inequalities of the form (1.2) and (1.3) is often referred to as “sprinkling”.

While **P1** – **P3** are general conditions on the family \mathbb{P}^u , $u \in (a, b)$, the next condition **S1** pertains to the connectivity properties of \mathcal{S} . It can be understood as a certain local uniqueness property for macroscopic connected components in \mathcal{S} . Roughly speaking, this condition implies that with high probability, large enough boxes intersect exactly one connected component of \mathcal{S} with large diameter.

$$\begin{aligned} &\text{For } r \in [0, \infty], \text{ we denote by } \mathcal{S}_r, \text{ the set of vertices of } \mathcal{S} \\ &\text{which are in connected components of } \mathcal{S} \text{ of } (\ell^1\text{-})\text{diameter } \geq r. \end{aligned} \quad (1.5)$$

In particular, in the case $r = \infty$ we recover the set \mathcal{S}_∞ of vertices of \mathcal{S} contained in infinite connected components of \mathcal{S} .

S1 There exists a function $f_s : (a, b) \times \mathbb{Z}_+ \rightarrow \mathbb{R}$ such that

$$\begin{aligned} &\text{for each } u \in (a, b), \text{ there exist } \Delta_s = \Delta_s(u) > 0 \text{ and } R_s = R_s(u) < \infty \\ &\text{such that } f_s(u, R) \geq (\log R)^{1+\Delta_s} \text{ for all } R \geq R_s, \end{aligned} \quad (1.6)$$

and for all $u \in (a, b)$ and $R \geq 1$, the following inequalities are satisfied:

$$\mathbb{P}^u[\mathcal{S}_R \cap B(R) \neq \emptyset] \geq 1 - e^{-f_s(u, R)},$$

and

$$\mathbb{P}^u \left[\bigcap_{x, y \in \mathcal{S}_{R/10} \cap B(R)} \{x \text{ is connected to } y \text{ in } \mathcal{S} \cap B(2R)\} \right] \geq 1 - e^{-f_s(u, R)}.$$

It is not difficult to see that if the family \mathbb{P}^u , $u \in (a, b)$, satisfies **S1**, then for any $u \in (a, b)$,

$$\mathbb{P}^u\text{-a.s.}, \text{ the set } \mathcal{S}_\infty \text{ is non-empty and connected,} \quad (1.7)$$

and there exist $c_1 = c_1(u) > 0$ and $C_1 = C_1(u) < \infty$ such that for all $R \geq 1$,

$$\mathbb{P}^u [\mathcal{S}_\infty \cap B(R) \neq \emptyset] \geq 1 - C_1 e^{-c_1 f_S(u, R)}. \quad (1.8)$$

Moreover, using a standard covering argument, if the family \mathbb{P}^u , $u \in (a, b)$, is invariant under the lattice shifts and satisfies **S1**, then for any $u \in (a, b)$ and $\varepsilon > 0$, there exist $c_2 = c_2(u, \varepsilon) > 0$ and $C_2 = C_2(u, \varepsilon) < \infty$ such that for all $R \geq 1$,

$$\mathbb{P}^u \left[\bigcap_{x, y \in \mathcal{S}_{\varepsilon R} \cap B(R)} \{x \text{ is connected to } y \text{ in } \mathcal{S} \cap B((1 + \varepsilon)R)\} \right] \geq 1 - C_2 e^{-c_2 f_S(u, R)}. \quad (1.9)$$

We omit the proof of (1.9) and refer the reader to the derivation of Proposition 1 from Lemma 13 in [21], where essentially the same statement is proved.

Our final condition **S2** on the measures \mathbb{P}^u , $u \in (a, b)$, concerns the density of \mathcal{S}_∞ ,

$$\eta(u) = \mathbb{P}^u [0 \in \mathcal{S}_\infty]. \quad (1.10)$$

S2 The function $\eta(\cdot)$ is positive and continuous on (a, b) .

In fact, the positivity of $\eta(\cdot)$ already follows from (1.7), if we assume that \mathbb{P}^u , $u \in (a, b)$, are invariant with respect to lattice shifts and satisfy **S1**. Note that if we assume **P2**, then

$$\eta(\cdot) \text{ is non-decreasing on } (a, b).$$

Remark 1.2. Condition **S2** is the price that we pay for asking for decorrelation inequalities only in the weak form of **P3** (see the proofs of Lemmas 4.2 and 4.4). If the inequalities (1.2) and (1.3) in **P3** were true for $\hat{u} = u$, then we would not need the assumption **S2**. However, as we discuss in Remark 1.1, there are many percolation models with long-range correlations for which (1.2) and (1.3) fail to hold with $\hat{u} = u$, but conditions **P3** and **S2** are still valid.

The main result of our paper is the following theorem. For $x, y \in \mathcal{S}$, let $\rho_S(x, y) \in \mathbb{Z}_+ \cup \{\infty\}$ denote the *chemical distance* (also known as the internal or graph distance) in \mathcal{S} between x and y , i.e.,

$$\rho_S(x, y) = \inf \left\{ n \geq 0 : \begin{array}{l} \text{there exist } x_0, \dots, x_n \in \mathcal{S} \text{ such that } x_0 = x, x_n = y, \\ \text{and } |x_k - x_{k-1}|_1 = 1 \text{ for all } k = 1, \dots, n \end{array} \right\}.$$

Here we use the usual convention $\inf \emptyset = \infty$, i.e., we set $\rho_S(x, y) = \infty$ if x and y are in different connected components of \mathcal{S} .

Theorem 1.3 (Chemical distance). *Assume that the family of probability measures \mathbb{P}^u , $u \in (a, b)$, on $(\{0, 1\}^{\mathbb{Z}^d}, \mathcal{F})$, $d \geq 2$, satisfies the conditions **P1** – **P3** and **S1** – **S2**. For any $u \in (a, b)$, there exist $c = c(u) > 0$ and $C = C(u) < \infty$ such that for all $R \geq 1$,*

$$\mathbb{P}^u [\text{for all } x, y \in \mathcal{S}_R \cap B(R), \rho_S(x, y) \leq CR] \geq 1 - C e^{-c(\log R)^{1+\Delta_S}}, \quad (1.11)$$

where Δ_S comes from condition **S1** (see (1.6)).

Remark 1.4. If a probability measure \mathbb{P}^u is invariant under the lattice shifts and satisfies **S1**, then a weak form of (1.11) holds. Namely, for any $\varepsilon > 0$, there exist $c = c(u, \varepsilon) > 0$ and $C = C(u, \varepsilon) < \infty$ such that for all $R \geq 1$,

$$\mathbb{P}^u \left[\text{for all } x, y \in \mathcal{S}_R \cap B(R), \rho_S(x, y) \leq CR^{1+\varepsilon} \right] \geq 1 - Ce^{-c(\log R)^{1+\Delta_S}}.$$

However, in order to prove (1.11), we need to quantify correlations in some way. In this paper we achieve this by assuming **P3**.

Using standard methods, we can deduce from Theorem 1.3 a shape theorem for ρ_S -balls of \mathcal{S}_∞ . For $x \in \mathcal{S}$ and $r \geq 0$, we denote by $B_S(x, r)$ the ball in \mathcal{S} with center x and radius $\lfloor r \rfloor$, i.e.,

$$B_S(x, r) = \{y \in \mathcal{S} : \rho_S(x, y) \leq r\}.$$

The shape theorem states that large balls in \mathcal{S}_∞ with respect to the metric ρ_S after rescaling have an asymptotic deterministic shape.

Theorem 1.5 (Shape theorem). *Assume that the family of probability measures \mathbb{P}^u , $u \in (a, b)$, satisfies **P1** – **P3** and **S1** – **S2**. Then for any $u \in (a, b)$ there exists a convex compact set $D_u = D_u(\mathbb{P}^u) \subset \mathbb{R}^d$ such that for each $\varepsilon \in (0, 1)$, there exists a $\mathbb{P}^u[\cdot | 0 \in \mathcal{S}_\infty]$ -almost surely finite random variable $\tilde{R}_{\varepsilon, u}$ satisfying*

$$\forall R \geq \tilde{R}_{\varepsilon, u} : \mathcal{S}_\infty \cap (1 - \varepsilon)R \cdot D_u \subseteq B_S(0, R) \subseteq \mathcal{S}_\infty \cap (1 + \varepsilon)R \cdot D_u.$$

Remark 1.6. (i) The set D_u preserves symmetries of \mathbb{P}^u . In particular, if \mathbb{P}^u is invariant with respect to the isometries of \mathbb{R}^d which preserve 0 and \mathbb{Z}^d , then so is D_u .

(ii) Since \mathbb{P}^u , $u \in (a, b)$, satisfy **P2**, it follows that for any $u, u' \in (a, b)$ with $u < u'$, $D_u \subseteq D_{u'}$.

We prove Theorem 1.5 in Section 8 using Theorem 1.3 and Kingman’s subadditive ergodic theorem [14] in a standard fashion.

Let us now describe the strategy for the proof of Theorem 1.3. We will set up a general multi-scale renormalization scheme, which involves the recursive definition of unfavorable regions on higher and higher scales. Our aim is to iteratively construct short nearest neighbor paths in \mathcal{S} : given a short path π (not necessarily in \mathcal{S}) which avoids the unfavorable regions on a given scale, we modify this path in a way that the resulting path π' (still not necessarily in \mathcal{S}) avoids the unfavorable regions on the previous scale as well. We choose our scales to grow faster than geometrically in order to achieve an upper bound on the length of the iteratively constructed path which is uniform in the number of iterations. Our definition of the favorable region on the bottom scale implies (due to **S1**) that such region contains a unique macroscopic connected component which is locally connected to the macroscopic connected components of all the neighboring favorable regions. By “glueing” such connected components together, we are able to locally modify short paths (not necessarily in \mathcal{S}) that avoid unfavorable regions on the bottom scale to obtain short nearest neighbor paths in \mathcal{S} . The weak decorrelation inequalities of **P3** allow us to set up a renormalization scheme in which unfavorable regions on higher scales become very unlikely. This is a tricky part, since we should be satisfied only with monotone events and cope with different parameters (u and \hat{u}) in the decorrelation inequalities. All in all, favorable regions on high scales are likely and allow for short paths in \mathcal{S} . We use this conclusion to define an event \mathcal{H} such that (a) \mathcal{H} implies the event on the left-hand side of (1.11) and (b) $\mathbb{P}^u[\mathcal{H}]$ is at least $1 - Ce^{-c(\log R)^{1+\Delta_S}}$.

Open problems. It is natural to ask whether the conditions **P1** – **P3** and **S1** – **S2** are sufficient for obtaining finer results about the structure of \mathcal{S}_∞ . More specifically, we ask whether the conditions **P1** – **P3** and **S1** – **S2** imply quenched *Gaussian bounds on the transition kernel* of a simple random walk on \mathcal{S}_∞ and a quenched *invariance principle* for the walk.

In the case of Bernoulli percolation, the corresponding Gaussian bounds were proved in [3] (see also [16] for a partial result) using extensively techniques and results of [2], and the quenched invariance principle was subsequently proved in [24] for $d \geq 4$ and in [5, 17] for all $d \geq 2$ using the results of [3]. Apart from Bernoulli percolation, the above questions are still open even for the specific examples treated in Sections 2.2–2.4. We should mention here though that a slightly subdiffusive upper bound on the on-diagonal terms of the transition kernel of a simple random walk on random interlacements (see Section 2.2) has been recently obtained in [18, Theorem 2.3].

We hope that the multi-scale renormalization scheme developed in this paper will find applications in the proof of the above questions not only for specific examples of models, but in the full generality of assumptions **P1** – **P3** and **S1** – **S2**.

The rest of the paper is organized as follows.

In Section 2 we give applications of our main results by demonstrating particular models in which the conditions **P1** – **P3** and **S1** – **S2** are satisfied. Our main focus is on the models for which *only* weak decorrelation inequalities of the form (1.2) and (1.3) are available. In particular, in Section 2.2 we show that our results give an alternative approach to the results of [8] for random interlacements, in Section 2.3 we show that our results hold for the vacant set of random interlacements in a (sub)phase of the supercritical parameters, and in Section 2.4 we show that the level sets of the Gaussian free field fit into our setup for a (sub)phase of the supercritical parameters. The results of Sections 2.3 and 2.4 are new. Finally, in Section 2.5 we apply results of Section 2.3 to obtain new results about the (chemical) diameter of the giant connected component in the complement of the random walk trace on the torus.

In Section 3 we define the multi-scale renormalization scheme in an abstract setting. We define the notion of unfavorable regions and state conditions under which the higher scale unfavorable regions are unlikely.

In Section 4 we define the connectivity patterns of favorable regions on the bottom scale of our scheme, bringing into play two families of events of high probability, one increasing and one decreasing.

In Section 5 we use the events from Section 4 and the notion of favorable/unfavorable regions from Section 3 to define good and bad regions on all scales. We then describe explicitly how to modify a path through a good region on one scale to obtain a path through a good region on a lower scale without increasing its length by too much. This gives us one iteration in the construction of short paths in \mathcal{S} .

In Section 6 we prove Theorem 1.3.

In Section 7 we prove the main result of Section 3 (see Theorem 3.1) verifying the conditions under which high scale regions are favorable with very high probability. We use the weak decorrelation inequalities to prove that a good upper bound on the probability of unfavorable events on a certain scale results in an even better bound on the next scale.

In Section 8 we prove Theorem 1.5 using Theorem 1.3 and Kingman’s subadditive ergodic theorem [14] in a standard fashion.

Throughout the paper, e_i , $1 \leq i \leq d$, denote the canonical basis in \mathbb{R}^d . Constants are denoted by c and C . Their values may change from place to place. We omit the dependence of constants on a , b , and d , but reflect the dependence on other parameters in the notation.

2 Applications

In this section we give a number of examples for which our conditions and results hold. In particular, we show that Theorems 1.3 and 1.5 hold for

- (a) the *random interlacements* at any level $u > 0$ (see Section 2.2), giving alternative, model independent proofs to some of the results in [8],
- (b) the *vacant set of random interlacements* at level u in the (non-empty) regime of so-called “local uniqueness”, which is believed to coincide with the whole supercritical phase (see Section 2.3),
- (c) the *level sets of the Gaussian free field*, also in the (non-empty) regime of local uniqueness (see Section 2.4).

The results in Sections 2.3 and 2.4 are new. Let us also point out that the above models (a), (b) and (c) give rise to random subsets of \mathbb{Z}^d with *polynomial decay of correlations*, but still they satisfy the strong concentration bound (1.11).

In Section 2.5 we apply the results of Section 2.3 to study the (chemical) diameter of the largest connected component in the complement of the trace of random walk on a discrete torus.

2.1 Bernoulli percolation

Bernoulli site percolation with parameter $u \in (p_c, 1)$ on \mathbb{Z}^d , $d \geq 2$, satisfies all the conditions **P1** – **P3** and **S1** – **S2** (see, e.g., [12]). In this case \mathbb{P}^u is just the product measure on $(\{0, 1\}^{\mathbb{Z}^d}, \mathcal{F})$ with $\mathbb{P}^u[\Psi_x = 1] = 1 - \mathbb{P}^u[\Psi_x = 0] = u$.

The analogue of Theorem 1.3 for the chemical distance on the infinite cluster of Bernoulli percolation in the whole supercritical phase was proved in [2, Theorem 1.1], where the authors obtained an exponentially decaying upper bound.

2.2 Random interlacements

Random interlacements \mathcal{I}^u at level $u > 0$ on \mathbb{Z}^d , $d \geq 3$, is a random subset of \mathbb{Z}^d , which arises as the local limit as $N \rightarrow \infty$ of the set of sites visited by a simple random walk on the discrete torus $(\mathbb{Z}/N\mathbb{Z})^d$, $d \geq 3$, when it runs up to time $\lfloor uN^d \rfloor$, see [26, 31]. The distribution of \mathcal{I}^u is determined by the equations

$$\mathbb{P}[\mathcal{I}^u \cap K = \emptyset] = e^{-u \cdot \text{cap}(K)}, \quad \text{for any finite } K \subseteq \mathbb{Z}^d, \quad (2.1)$$

where $\text{cap}(K)$ denotes the discrete capacity of K . For any $u > 0$, \mathcal{I}^u is an infinite almost surely connected random subset of \mathbb{Z}^d (see [26, (2.21)]) with polynomially decaying correlations

$$|\mathbb{P}[x, y \in \mathcal{I}^u] - \mathbb{P}[x \in \mathcal{I}^u] \cdot \mathbb{P}[y \in \mathcal{I}^u]| \asymp (1 + |x - y|)^{2-d}, \quad x, y \in \mathbb{Z}^d \quad (2.2)$$

(see [26, (1.68)]). Geometric properties of random interlacements have been extensively studied in the last few years (see [8, 15, 18, 19, 20, 21, 22]). In particular, the fact that the conclusions of Theorems 1.3 and 1.5 hold for the random interlacements \mathcal{I}^u at level $u > 0$ has been recently established in [8, Theorems 1.3 and 1.1]. The key idea in the proofs of [8] is a certain refinement of the strategy developed in [20], which crucially relies on the underlying random walk structure of random interlacements (see [26, (1.53)]). The goal of this section is to observe (based on earlier results about random interlacements) that the distributions of random interlacements at level u satisfy conditions **P1** – **P3** and **S1** – **S2**. As a result, we obtain an alternative, model independent approach to the results of [8].

Theorem 2.1. *For any $a > 0$, the distributions of \mathcal{I}^u , $u \in (a, \infty)$, satisfy conditions **P1** – **P3** and **S1** – **S2**. In particular, the results of Theorems 1.3 and 1.5 hold for \mathbb{P}^u being the distribution of \mathcal{I}^u for any $u > 0$.*

Remark 2.2. Theorem 1.3 applied to the distribution of \mathcal{I}^u gives a weaker result than [8, Theorem 1.3], since the bound on the right-hand side of (1.11) is not as good as the stretched exponential bound in [8, Theorem 1.3]. Nevertheless, this weaker bound is still sufficient for us to deduce Theorem 1.5, thus giving an alternative proof of [8, Theorem 1.1].

Proof of Theorem 2.1. The distributions of \mathcal{I}^u , $u > 0$, satisfy condition **P1** by [26, (2.3)], and condition **P2** by [26, (1.53)].

The fact that the distributions of \mathcal{I}^u , $u \in (a, \infty)$, satisfy condition **P3** for any given $a > 0$ follows from the decoupling inequalities [27, Theorem 2.1] applied to the graph $E = \mathbb{Z}^{d-1} \times \mathbb{Z}$ (see also [27, Remark 2.7(1)]). We take $\chi_P = 1/4$ in (1.1). Our choice of parameters in [27, Theorem 2.1] is the following:

$$K = 1, \quad n = 0, \quad L_0 = L, \quad \ell_0 = \lfloor |x_1 - x_2|/L \rfloor, \quad \nu = d - 2 \quad (\alpha = d - 1, \beta = 2), \quad \nu' = \nu/4.$$

We also use the notation \hat{u} instead of u' here. Note that with the above choice of parameters we have $\nu \geq 1$ and $\ell_0 \geq R$ (since $d \geq 3$ and $|x_1 - x_2| \geq R \cdot L$). In [27, Theorem 2.1], the inequality $\ell_0 \geq c(K, \nu')$ is required, so we need to choose $R_P \geq c(1, \frac{d-2}{4})$. With the above choice of parameters, u and \hat{u} satisfy condition (2.7) of [27, Theorem 2.1] for all $R \geq R_P$ (for some suitably big R_P), so (1.2) and (1.3) immediately follow from [27, (2.8) and (2.9)] with the function $f_P(L) = 2aL - \ln 2$.

By [26, (2.21)], for any $u > 0$, \mathcal{I}^u is an almost surely connected infinite subset of \mathbb{Z}^d , in particular we have $\mathcal{I}_r^u = \mathcal{I}^u$ (see (1.5)) for any $r \in [0, \infty]$. The distributions of \mathcal{I}^u , $u > 0$, satisfy **S1** by [21, Proposition 1], and **S2** by (2.1) (since $\mathbb{P}[0 \in \mathcal{I}^u] = 1 - e^{-u \text{cap}(\{0\})}$ is a positive and continuous function of $u \in (0, \infty)$).

We have checked that the distributions of \mathcal{I}^u satisfy **P1** – **P2** and **S1** – **S2** for all $u > 0$, and **P3** for $u > a > 0$ (and any given $a > 0$). The proof of Theorem 2.1 is complete. \square

2.3 Vacant set of random interlacements

The vacant set of random interlacements \mathcal{V}^u at level u in dimension $d \geq 3$ is defined as the complement of random interlacements \mathcal{I}^u at level u :

$$\mathcal{V}^u = \mathbb{Z}^d \setminus \mathcal{I}^u, \quad u > 0. \quad (2.3)$$

In particular, it follows from (2.1) that the distribution of \mathcal{V}^u is determined by the equations

$$\mathbb{P}[\mathcal{V}^u \supset K] = e^{-u \cdot \text{cap}(K)}, \quad \text{for any finite } K \subseteq \mathbb{Z}^d,$$

and it follows from (2.2) that the correlations in \mathcal{V}^u decay polynomially.

In the same way as random interlacements, the vacant set at level $u > 0$ arises as the local limit as $N \rightarrow \infty$ of the set of sites not visited by a simple random walk (we call it the vacant set, too) on the discrete torus $(\mathbb{Z}/N\mathbb{Z})^d$, $d \geq 3$, when it runs up to time $\lfloor uN^d \rfloor$. In fact, the connection between the vacant sets of random interlacements and a simple random walk on $(\mathbb{Z}/N\mathbb{Z})^d$ is in terms of a strong coupling of [30, Theorem 1.1]. This coupling provided a powerful tool to study the fragmentation of $(\mathbb{Z}/N\mathbb{Z})^d$ by a simple random walk using existing results about the geometry of \mathcal{V}^u . In particular, it was shown in [30] that the fragmentation question is intimately related to the existence of a non-trivial phase transition in u for \mathcal{V}^u : there exists $u_* \in (0, \infty)$ such that

- (i) for any $u > u_*$, almost surely, all connected components of \mathcal{V}^u are finite, and
- (ii) for any $u < u_*$, almost surely, \mathcal{V}^u contains an infinite connected component.

The fact that $u_* < \infty$ was proved in [26], and the positivity of u_* was established in [26] when $d \geq 7$, and later in [25] for all $d \geq 3$. It is also known (see [28]) that if there exists an infinite connected component in \mathcal{V}^u , then it is almost surely unique.

In this section we focus on the supercritical phase of \mathcal{V}^u (regime (ii) above). More specifically, we would like to derive conclusions of Theorems 1.3 and 1.5 for the vacant set \mathcal{V}^u . While it would be desirable to obtain results for all $u < u_*$, our understanding of the supercritical phase is not good enough to do that. In Theorem 2.3 we prove that there exists $\bar{u} \in (0, u_*]$ such that the conclusions of Theorems 1.3 and 1.5 hold for \mathcal{V}^u with $u \in (0, \bar{u})$. We believe that $\bar{u} = u_*$ (see Remark 2.4). Even in the present form, the result of Theorem 2.3 below is new for any choice of dimension $d \geq 3$ and the level $u > 0$.

Theorem 2.3. *There exists $\bar{u} \in (0, u_*]$ such that for any $1/\bar{u} < b < \infty$, the distributions \mathbb{P}^v of $\mathcal{V}^{1/v}$, $v \in (1/\bar{u}, b)$, satisfy conditions **P1** – **P3** and **S1** – **S2**. In particular, the conclusions of Theorems 1.3 and 1.5 hold for the law of \mathcal{V}^u for any $u \in (0, \bar{u})$.*

Proof of Theorem 2.3. For $v \in (0, \infty)$, let \mathbb{P}^v denote the distribution of $\mathcal{V}^{1/v}$. Recall from (2.3) that for any $v > 0$, $\mathcal{V}^{1/v} = \mathbb{Z}^d \setminus \mathcal{I}^{1/v}$. From this and Theorem 2.1 it is immediate that the family \mathbb{P}^v , $v \in (0, \infty)$, satisfies conditions **P1** and **P2**, and the family \mathbb{P}^v , $v \in (0, b)$, satisfies condition **P3** for any $b < \infty$.

Let \bar{u} be the largest u such that the family \mathbb{P}^v , $v \in (1/u, \infty)$, satisfies condition **S1**. It follows from (1.7) that $\bar{u} \leq u_*$. The positivity of \bar{u} for all $d \geq 3$ follows from the main result of [9] (for $d \geq 5$, it also follows from [29, (1.2) and (1.3)]).

Finally, the family \mathbb{P}^v , $v \in (1/u_*, \infty)$, satisfies condition **S2** by [28, Corollary 1.2].

We have checked that the distributions of $\mathcal{V}^{1/v}$ satisfy **P1** – **P2** for $v \in (0, \infty)$, **P3** for $v \in (0, b)$ (and any given $b < \infty$), **S1** for $v \in (1/\bar{u}, \infty)$ (with $\bar{u} > 0$), and **S2** for $v \in (1/u_*, \infty)$. Since $\bar{u} \leq u_*$, the proof of Theorem 2.3 is complete. \square

Remark 2.4. Assumption **S1** is satisfied by Bernoulli percolation in the whole supercritical phase (see e.g. [12, (7.89),(8.98)]). Thus it is reasonable to conjecture that $\bar{u} = u_*$, i.e., that the distributions of \mathcal{V}^u satisfy **S1** for all $u < u_*$. The verification of this conjecture would in particular imply that the conclusions of Theorems 1.3 and 1.5 hold for \mathcal{V}^u for any $u \in (0, u_*)$.

2.4 Level sets of the Gaussian free field

The Gaussian free field on \mathbb{Z}^d , $d \geq 3$, is a centered Gaussian field $\varphi = (\varphi_x)_{x \in \mathbb{Z}^d}$ under the probability measure \mathbb{P} with covariance $\mathbb{E}[\varphi_x \varphi_y] = g(x, y)$, for all $x, y \in \mathbb{Z}^d$, where $g(\cdot, \cdot)$ denotes the Green function of the simple random walk on \mathbb{Z}^d . The random field φ exhibits long-range correlations, since $g(0, x)$ decays like $|x|^{2-d}$ as $|x| \rightarrow \infty$.

For any $h \in \mathbb{R}$, we define the *excursion set* above level h as

$$E_\varphi^{\geq h} = \{x \in \mathbb{Z}^d : \varphi_x \geq h\}. \quad (2.4)$$

We view $E_\varphi^{\geq h}$ as a random subgraph of \mathbb{Z}^d . For any $d \geq 3$, there exists $h_* \in [0, \infty)$ such that

- (a) for any $h < h_*$, \mathbb{P} -almost surely $E_\varphi^{\geq h}$ contains a unique infinite connected component, and
- (b) for any $h > h_*$, \mathbb{P} -almost surely all the connected components of $E_\varphi^{\geq h}$ are finite.

The finiteness of h_* was established in [6] when $d = 3$, and later in [23] for all $d \geq 3$. The non-negativity of h_* was shown in [6], and the uniqueness in [23, Remark 1.6].

In this section we show that the excursion set $E_\varphi^{\geq h}$ of the Gaussian free field satisfies the assumptions and results of this paper in a certain sub-regime $(-\infty, \bar{h})$ of the supercritical phase $(-\infty, h_*)$. The relation between \bar{h} and h_* is discussed in Remark 2.7.

The main result of this section is the following theorem.

Theorem 2.5. *For $u > 0$, let $h(u) = h_* - u + 1$, and let \mathbb{P}^u be the law of $E_\varphi^{\geq h(u)}$. Then for any $d \geq 3$, there exists $\bar{u} \in [1, \infty)$ such that the family of distributions \mathbb{P}^u , $u > \bar{u}$, satisfies assumptions **P1** – **P3** and **S1** – **S2**. In particular, there exists $\bar{h} \in (-\infty, h_*]$ such that the conclusions of Theorems 1.3 and 1.5 hold for the distribution of the excursion set $E_\varphi^{\geq h}$ for any $h \in (-\infty, \bar{h})$.*

Proof. The assumption **P1** is satisfied by \mathbb{P}^u , $u > 0$, see the discussion above [23, Lemma 1.5].

The family \mathbb{P}^u , $u > 0$, satisfies **P2**, since the inclusion $E_\varphi^{\geq h} \subseteq E_\varphi^{\geq h'}$ holds for all $h' \leq h \in \mathbb{R}$ by (2.4).

The fact that the distributions \mathbb{P}^u , $u \in (1, \infty)$, satisfy **P3** follows from Lemma 2.6 below.

Recall that \mathcal{F} denotes the canonical sigma-algebra on $\{0, 1\}^{\mathbb{Z}^d}$, and denote by \mathcal{G} the canonical sigma-algebra on $\mathbb{R}^{\mathbb{Z}^d}$. For an event A in \mathcal{F} or \mathcal{G} , we denote by $\mathbf{1}(A)$ the indicator of A . For any $B \in \mathcal{F}$ and $h \in \mathbb{R}$, we define the event $B^h \in \mathcal{G}$ by

$$\forall \varphi \in \mathbb{R}^{\mathbb{Z}^d} : \quad \varphi \in B^h \iff \left(\mathbf{1}(\varphi_x \geq h), x \in \mathbb{Z}^d \right) \in B. \quad (2.5)$$

Lemma 2.6. *Let $d \geq 3$. Let $L \geq 1$ and $R \geq 50$ be integers. Let $x_1, x_2 \in \mathbb{Z}^d$. For $i \in \{1, 2\}$, let $A_i \in \sigma(\Psi_y : y \in B(x_i, 10L))$ be decreasing events, and $B_i \in \sigma(\Psi_y : y \in B(x_i, 10L))$ increasing events. Let $h, \hat{h} \in \mathbb{R}$. There exist $c > 0$ and $C < \infty$ such that if*

$$\hat{h} \geq h + C \cdot R^{2-d} \quad \text{and} \quad |x_1 - x_2|_\infty \geq R \cdot L,$$

then

$$\mathbb{P} \left[A_1^h \cap A_2^h \right] \leq \mathbb{P} \left[A_1^{\hat{h}} \right] \cdot \mathbb{P} \left[A_2^{\hat{h}} \right] + C e^{-L^c} \quad (2.6)$$

and

$$\mathbb{P} \left[B_1^{\hat{h}} \cap B_2^{\hat{h}} \right] \leq \mathbb{P} \left[B_1^h \right] \cdot \mathbb{P}^u \left[B_2^h \right] + C e^{-L^c}. \quad (2.7)$$

The proof of Lemma 2.6 is similar to that of [23, Proposition 2.2], but some new ideas are needed in order to overcome the issue discussed in [23, Remark 2.3 (3)]. We postpone the proof of Lemma 2.6 to after finishing the rest of the proof of Theorem 2.5.

To see that Lemma 2.6 implies that **P3** is satisfied by the family \mathbb{P}^u , $u \in (1, \infty)$, note that for any $u, \hat{u} > 1$ such that $u \geq (1 + C \cdot R^{2-d})\hat{u}$, we have that $u \geq \hat{u} + C \cdot R^{2-d}$. In particular, we get that $h(\hat{u}) = h_* - \hat{u} + 1 \geq h(u) + C \cdot R^{2-d}$. It is now immediate from Lemma 2.6 that \mathbb{P}^u , $u \in (1, \infty)$, satisfies **P3** for any choice of $0 < \varepsilon_P < 1$ and $0 < \chi_P < d - 2$.

We proceed with **S1**. Let $\bar{u} \in [0, \infty]$ be the smallest u' such that the family \mathbb{P}^u , $u > u'$, satisfies condition **S1**. Define $\bar{h} = h(\bar{u})$ and note that $\bar{h} \leq h_*$ (i.e., $\bar{u} \geq 1$) follows from (1.7) and the definition of h_* . We will now prove that $\bar{u} < \infty$ (i.e., $\bar{h} > -\infty$).

We say that $x, y \in \mathbb{Z}^d$ are $*$ -connected in \mathcal{S} if there exists $x_1, \dots, x_n \in \mathcal{S}$ such that $x_1 = x$, $x_n = y$ and $|x_{k+1} - x_k|_\infty = 1$ for all $k \in \{1, \dots, n-1\}$. Essentially the same proof as the one of [23, (2.64)] implies that for any $d \geq 3$ there exists $\tilde{h} \in (0, \infty)$ and constants $c > 0$ and $C < \infty$ such that

$$\mathbb{P} \left[0 \text{ and } x \text{ are } * \text{-connected in } E_\varphi^{\geq h} \right] \leq C e^{-|x|^c}, \quad \text{for all } x \in \mathbb{Z}^d, h \geq \tilde{h}.$$

Observing that the laws of $E_\varphi^{\geq h}$ and $\mathbb{Z}^d \setminus E_\varphi^{\geq -h}$ are the same under \mathbb{P} , we immediately obtain that the $*$ -connected components in the complement of $E_\varphi^{\geq -h}$ are very small for any $h \in (\tilde{h}, \infty)$. Then, a standard argument based on the connectedness of $*$ -boundaries (see, e.g., the proof of [9, Corollary 3.7]) implies that the distribution of $E_\varphi^{\geq -h}$ satisfies **S1** for any $h \in (\tilde{h}, \infty)$. Let $\tilde{u} = h_* + \tilde{h} + 1 < \infty$. For any $u > \tilde{u}$, we have that $h(u) < -\tilde{h}$. Therefore, the family \mathbb{P}^u , $u > \tilde{u}$ satisfies condition **S1**. In particular, $\bar{u} \leq \tilde{u} < \infty$.

Finally, recall that by [23, Remark 1.6], for any $h < h_*$, \mathbb{P} -almost surely, there is a unique infinite connected component in $E_\varphi^{\geq h}$. Therefore, with standard methods (see, e.g., the proof of [12, Lemma (8.10)]), it follows that the distributions of \mathbb{P}^u , $u \in (1, \infty)$, satisfy assumption **S2**.

We have checked that the family of distributions \mathbb{P}^u satisfies **P1** – **P2** for $u > 0$, **P3** and **S2** for $u \in (1, \infty)$, and **S1** for $u \in (\bar{u}, \infty)$ with $\bar{u} < \infty$. The proof of Theorem 2.5 is complete, given the result of Lemma 2.6. \square

Remark 2.7. Similarly to Remark 2.4, it is reasonable to conjecture that $\bar{h} = h_*$, i.e., that the distribution of $E_\varphi^{\geq h}$ satisfies **S1** for all $h < h_*$. As soon as this conjecture is verified, the conclusions of Theorems 1.3 and 1.5 extend to the excursion set $E_\varphi^{\geq h}$ for any parameter h in the supercritical regime $(-\infty, h_*)$. Proving that $\bar{h} = h_*$ may be quite hard. A reasonable start would be to prove that $\bar{h} \geq 0$. A strict inequality, i.e., $\bar{h} > 0$, would be even better, but at the moment there is no rigorous argument even for the statement that $h_* > 0$ for all $d \geq 3$. It is only known that $h_* > 0$ when the dimension d is large, see [23, Theorem 3.3]. Extending the result of [23, Theorem 3.3] to all dimensions $d \geq 3$ is thus another interesting open problem, see [23, Remark 3.6 (3)].

Proof of Lemma 2.6. We will only prove (2.7). The proof of (2.6) is similar, and we omit it.

Let $K_i = B(x_i, 10L)$ for $i = 1, 2$. Let $H(h)$ be the event that

$$H(h) = \left\{ \text{there is no nearest neighbour path in } E_\varphi^{\geq h} \text{ connecting } K_1 \text{ to } B(x_1, 20L)^c \right\}.$$

By [23, (2.63)], there exists $h_0 \in (h_*, \infty)$ and constants $c > 0$ and $C < \infty$ such that $\mathbb{P}[H(h_0)^c] \leq Ce^{-L^c}$ for all $L \geq 1$. From now on we fix such h_0 and write H for $H(h_0)$. In order to prove (2.7) we only need to show that there exists a constant $C < \infty$ such that with γ defined as

$$\gamma = C \cdot R^{2-d}, \tag{2.8}$$

for every $h \in \mathbb{R}$ and every pair $B_i \in \sigma(\Psi_y : y \in K_i)$, $i = 1, 2$ of increasing events in \mathcal{F} satisfying the assumptions of Lemma 2.6, we have

$$\mathbb{P} \left[B_1^h \cap B_2^h \cap H \right] \leq \mathbb{P} \left[B_1^{h-\gamma} \right] \cdot \mathbb{P} \left[B_2^{h-\gamma} \right]. \tag{2.9}$$

(Here, to simplify notation, we replaced \hat{h} of (2.7) by h , and h of (2.7) by $h - \gamma$, hopefully without confusion.)

Denote by \mathcal{D} the union of K_1 and the set of vertices of \mathbb{Z}^d connected to K_1 by a simple path in $E_\varphi^{\geq h_0}$. Note that if H occurs, then

$$B(x_1, 10L) \subseteq \mathcal{D} \subseteq B(x_1, 20L).$$

Denote by \mathbf{D} the set of subsets D of $B(x_1, 20L)$ that can arise as a realization of \mathcal{D} with positive \mathbb{P} -probability. We start proving (2.9) by writing

$$\mathbb{P} \left[B_1^h \cap B_2^h \cap H \right] = \sum_{D \in \mathbf{D}} \mathbb{P} \left[B_1^h \cap B_2^h \cap \{\mathcal{D} = D\} \right]. \tag{2.10}$$

Note that the event $\{\mathcal{D} = D\}$ is measurable with respect to the sub-sigma-algebra of \mathcal{G} generated by $\varphi_{\overline{D}} = (\varphi_x)_{x \in \overline{D}}$, where $\overline{D} = \{x \in \mathbb{Z}^d : \exists y \in D : |x - y|_1 \leq 1\}$ denotes the l_1 -closure of D . Thus for any $D \in \mathbf{D}$,

$$\mathbb{P} \left[B_1^h \cap B_2^h \cap \{\mathcal{D} = D\} \right] = \mathbb{E} \left[\mathbf{1}(B_1^h) \cdot \mathbf{1}(\mathcal{D} = D) \cdot \mathbb{P} \left[(\varphi_x)_{x \in K_2} \in B_2^h \mid \varphi_{\overline{D}} \right] \right]. \quad (2.11)$$

According to [23, Remark 1.3], for every $D \in \mathbf{D}$ there exists a probability measure $\tilde{\mathbb{P}}$ on $\mathbb{R}^{\mathbb{Z}^d}$ such that $(\tilde{\varphi}_x)_{x \in \mathbb{Z}^d}$ is a centered Gaussian field under $\tilde{\mathbb{P}}$ with $\tilde{\varphi}_x = 0$ for all $x \in \overline{D}$, and

$$\mathbb{P} \left[(\varphi_x)_{x \in \mathbb{Z}^d} \in \cdot \mid \varphi_{\overline{D}} \right] = \tilde{\mathbb{P}} \left[(\tilde{\varphi}_x + \mu_x)_{x \in \mathbb{Z}^d} \in \cdot \right], \quad (2.12)$$

where μ_x , $x \in \mathbb{Z}^d$ is given by

$$\mu_x = \sum_{y \in \overline{D}} P_x \left[H_{\overline{D}} < \infty, X_{H_{\overline{D}}} = y \right] \cdot \varphi_y,$$

where P_x denotes the law of simple random walk on \mathbb{Z}^d started at x , $H_{\overline{D}}$ denotes the first time that the random walk enters \overline{D} and $X_{H_{\overline{D}}}$ is the position of the walker at time $H_{\overline{D}}$.

Let us fix $D \in \mathbf{D}$ and $\varphi_{\overline{D}}$ such that the event $\{\mathcal{D} = D\}$ occurs. By the definition of \mathcal{D} , we have that $\mu_x < h_0 \cdot P_x(H_{\overline{D}} < \infty)$ for all $x \notin \overline{D}$. Under the assumptions of Lemma 2.6 the l^∞ -distance of K_2 and \overline{D} is at least $\frac{1}{5}R \cdot L$, thus by a standard argument based on estimates for the discrete Green function and capacity (see, e.g., the calculation below [23, (2.30)]), we obtain that there exists $C' < \infty$ such that for any $x \in K_2$, $P_x(H_{\overline{D}} < \infty) \leq C' \cdot R^{2-d}$. In particular, by taking C in (2.8) to be $2h_0 \cdot C'$, we obtain that

$$\max_{x \in K_2} \mu_x \leq h_0 \cdot C' \cdot R^{2-d} \stackrel{(2.8)}{=} \frac{1}{2}\gamma. \quad (2.13)$$

Now we are ready to carry out the “sprinkling” mentioned in Remark 1.1:

$$\begin{aligned} \mathbb{P} \left[(\varphi_x)_{x \in K_2} \in B_2^h \mid \varphi_{\overline{D}} \right] &\stackrel{(2.12)}{=} \tilde{\mathbb{P}} \left[(\tilde{\varphi}_x + \mu_x)_{x \in K_2} \in B_2^h \right] \stackrel{(*)}{\leq} \\ &\tilde{\mathbb{P}} \left[\left(\tilde{\varphi}_x + \frac{1}{2}\gamma \right)_{x \in K_2} \in B_2^h \right] = \tilde{\mathbb{P}} \left[\left(\tilde{\varphi}_x - \frac{1}{2}\gamma \right)_{x \in K_2} \in B_2^{h-\gamma} \right] \stackrel{(*)}{\leq} \tilde{\mathbb{P}} \left[(\tilde{\varphi}_x - \mu_x)_{x \in K_2} \in B_2^{h-\gamma} \right], \end{aligned} \quad (2.14)$$

where in the equations marked by $(*)$ above we used (2.13) and the fact that B_2 is an increasing event.

For any $B \in \mathcal{F}$ we define the *flipped* event $\check{B} \in \mathcal{F}$ by

$$\forall \xi \in \{0, 1\}^{\mathbb{Z}^d} : \quad \xi \in \check{B} \iff \left(\mathbf{1}(\xi_x = 0), x \in \mathbb{Z}^d \right) \in B. \quad (2.15)$$

Note that if B is an increasing event, then \check{B} is a decreasing event. With definition (2.15) at hand, we note that

$$\begin{aligned} \left\{ (\tilde{\varphi}_x - \mu_x)_{x \in K_2} \in B_2^{h-\gamma} \right\} &\stackrel{(2.5)}{=} \left\{ (\mathbf{1}(\tilde{\varphi}_x - \mu_x \geq h - \gamma), x \in K_2) \in B_2 \right\} \stackrel{(2.15)}{=} \\ &\left\{ (\mathbf{1}(\tilde{\varphi}_x - \mu_x < h - \gamma), x \in K_2) \in \check{B}_2 \right\} \stackrel{(2.5)}{=} \left\{ (-\tilde{\varphi}_x + \mu_x)_{x \in K_2} \in \check{B}_2^{\gamma-h} \right\}, \quad \tilde{\mathbb{P}}\text{-a.s.} \end{aligned} \quad (2.16)$$

Plugging this identity into (2.14) we obtain

$$\mathbb{P} \left[B_2^h \mid \varphi_{\overline{D}} \right] \leq \tilde{\mathbb{P}} \left[(-\tilde{\varphi}_x + \mu_x)_{x \in K_2} \in \check{B}_2^{\gamma-h} \right] \stackrel{(*)}{=} \tilde{\mathbb{P}} \left[(\tilde{\varphi}_x + \mu_x)_{x \in K_2} \in \check{B}_2^{\gamma-h} \right] \stackrel{(2.12)}{=} \mathbb{P} \left[\check{B}_2^{\gamma-h} \mid \varphi_{\overline{D}} \right],$$

where in the equation marked by (*) we used the fact that $\tilde{\varphi}_{K_2}$ and $-\tilde{\varphi}_{K_2}$ have the same distribution under $\tilde{\mathbb{P}}$. Substituting this inequality back into (2.11), we obtain

$$\mathbb{P} \left[B_1^h \cap B_2^h \cap \{\mathcal{D} = D\} \right] \leq \mathbb{P} \left[B_1^h \cap \check{B}_2^{\gamma-h} \cap \{\mathcal{D} = D\} \right].$$

Combining this inequality with (2.10) we conclude

$$\mathbb{P} \left[B_1^h \cap B_2^h \cap H \right] \leq \mathbb{P} \left[B_1^h \cap \check{B}_2^{\gamma-h} \right] \stackrel{(a)}{\leq} \mathbb{P} \left[B_1^h \right] \mathbb{P} \left[\check{B}_2^{\gamma-h} \right] \stackrel{(b)}{\leq} \mathbb{P} \left[B_1^{h-\gamma} \right] \mathbb{P} \left[B_2^{h-\gamma} \right],$$

where in (a) we used the fact that the increasing event B_1^h and the decreasing event $\check{B}_2^{\gamma-h}$ are negatively correlated under \mathbb{P} by the FKG-inequality for the Gaussian free field (see the remark above [23, Lemma 1.4]) and in (b) we used the fact that B_1 is increasing and a similar identity as in (2.16) combined with the fact that φ and $-\varphi$ have the same distribution under \mathbb{P} . This concludes the proof of (2.9) and (2.7). The proof of (2.6) is analogous and we omit it. The proof of Lemma 2.6 is complete. \square

2.5 Vacant set of random walk on a torus

We consider a simple symmetric nearest neighbor random walk on the d -dimensional torus $\mathbb{T}_N = (\mathbb{Z}/N\mathbb{Z})^d$, $d \geq 3$, started from a uniformly distributed vertex. Let $\mathcal{I}^{u,N}$ denote the random subset of \mathbb{T}_N which consists of the sites visited by the random walk in the first $\lfloor uN^d \rfloor$ steps, and

$$\mathcal{V}^{u,N} = \mathbb{T}_N \setminus \mathcal{I}^{u,N}$$

the vacant set of the random walk. We denote the distribution of $\mathcal{V}^{u,N}$ by $\mathbb{P}^{u,N}$. We view $\mathcal{V}^{u,N}$ as a (random) graph by drawing an edge between any two vertices of $\mathcal{V}^{u,N}$ at ℓ^1 -distance 1. The study of percolative properties of $\mathcal{V}^{u,N}$ was initiated in [4] and recently significantly boosted in [30]. Informally, [30, Theorems 1.2 and 1.3] state that for all $d \geq 3$, with high probability as $N \rightarrow \infty$,

- (a) if u is big enough then $\mathcal{V}^{u,N}$ consists of small connected components, but
- (b) if u is small enough then $\mathcal{V}^{u,N}$ contains macroscopic connected components.

It is conjectured that the transition between these two phases is sharp and occurs at the critical threshold u_* of the vacant set of random interlacements (see Section 2.3 for the definition of u_*).

Uniqueness of the giant component of $\mathcal{V}^{u,N}$ is only known in high dimensions: it follows from [30, Theorem 1.4] that if $d \geq 5$ and u is small enough then the connected component \mathcal{C}_{max}^u of $\mathcal{V}^{u,N}$ with the largest volume has a positive asymptotic density and the volume of the second largest connected component of $\mathcal{V}^{u,N}$ is bounded by $(\log N)^C$, with high probability as $N \rightarrow \infty$.

We are interested in the chemical distance $\rho_{u,N}(\cdot, \cdot)$ on the vacant set $\mathcal{V}^{u,N}$. In particular, we show in Theorem 2.8 below that if $d \geq 5$ and u is small enough then the (chemical) diameter of the subgraph of \mathbb{T}_N spanned by the giant vacant component \mathcal{C}_{max}^u is comparable to N with high probability as $N \rightarrow \infty$. Note that the analogous result about the chemical distances on $\mathcal{I}^{u,N}$ was proved in [8, Theorem 1.6] for all $d \geq 3$ and $u > 0$, which improved the earlier result of [18, Theorem 2.1].

Theorem 2.8. *For any $d \geq 5$, there exists $\tilde{u} = \tilde{u}(d) > 0$ such that for each $u < \tilde{u}$ there exists $C = C(d, u) < \infty$ such that for all $\kappa > 0$,*

$$\lim_{N \rightarrow \infty} N^\kappa \left(1 - \mathbb{P}^{u,N} \left[N/C \leq \max_{x,y \in \mathcal{C}_{max}^u} \rho_{u,N}(x,y) \leq CN \right] \right) = 0. \quad (2.17)$$

Remark 2.9. We believe that the statement of Theorem 2.8 holds for any $d \geq 3$ and $u < u_*$ (see Section 2.3). A possibly simpler open problem is to show that $\tilde{u}(d) > 0$ for $d \in \{3, 4\}$.

The main ingredients for the proof of Theorem 2.8 are Theorem 2.3, the strong coupling of [30, Theorem 1.1] between the vacant set of random walk on the torus and the vacant set of random interlacements, and the notion of *strongly supercritical* values of u in [30, Definition 2.4].

We split the proof of Theorem 2.8 into several steps. We first show that $\mathcal{V}^{u,N}$ contains many mesoscopic (not necessarily connected) subsets \mathcal{C}_z , $z \in \mathbb{T}_N$, with good chemical distance properties. We then show that these sets are connected to each other and to any vertex of $\mathcal{V}^{u,N}$ which is in a connected component of large enough diameter by paths of length at most CN . Most of the sets \mathcal{C}_z will be at distance $> N/C$ from each other, which will guarantee the lower bound on the (chemical) diameter of \mathcal{C}_{\max}^u .

Recall the definition of the vacant set of random interlacements \mathcal{V}^u from Section 2.3. We begin with a lemma, which is motivated by the strong coupling [30, Theorem 1.1] of $\mathcal{V}^{u,N}$ and \mathcal{V}^u . The lemma states that under some assumptions on u , there exists a mesoscopic subset of $\mathcal{V}^{u,N}$ with good chemical distance properties.

Let us choose (somewhat arbitrarily) the mesoscopic scale

$$n = N^{1/3}.$$

Lemma 2.10. *Let $d \geq 3$, $u > 0$, and $\varepsilon > 0$. If the distribution of $\mathcal{V}^{(1+\varepsilon)u}$ satisfies (1.8) and (1.11), then there exists $C = C(u, \varepsilon) < \infty$ such that for any $\kappa > 0$,*

$$\lim_{N \rightarrow \infty} N^\kappa \left(1 - \mathbb{P}^{u,N} \left[\begin{array}{l} \text{there exists } \mathcal{C} \subseteq \mathcal{V}^{u,N} \cap B(0, n) \text{ such that} \\ (i) \text{ for all } x, y \in \mathcal{C}, \rho_{u,N}(x, y) \leq Cn, \text{ and} \\ (ii) \text{ for all } x \in B(0, n/2), B(x, n^{1/d}) \cap \mathcal{C} \neq \emptyset \end{array} \right] \right) = 0. \quad (2.18)$$

Remark 2.11. Note that for any $u < \bar{u}$ (see the statement of Theorem 2.3), \mathcal{V}^u satisfies (1.8) and (1.11). In particular, we deduce from Lemma 2.10 and Theorem 2.3 that for any $d \geq 3$, $u < \bar{u}$, and $\kappa > 0$, (2.18) holds (by taking $\varepsilon > 0$ such that $(1 + \varepsilon)u < \bar{u}$).

Proof of Lemma 2.10. Let $A = B(0, n^2) \subset \mathbb{T}_N$ be the ball of radius $n^2 (= N^{2/3})$ centered at $0 \in \mathbb{T}_N$. The ball A is isomorphic to the ball $\mathbb{A} = B(0, n^2) \subset \mathbb{Z}^d$ via the graph isomorphism $\Phi_N : A \rightarrow \mathbb{A}$. It follows from [30, Theorem 1.1] that for any $d \geq 3$, $u > 0$, $\varepsilon > 0$, $\kappa > 0$, and $N \geq 1$, there exists a coupling $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mathbb{P}}^{u,N})$ of $\mathcal{V}^{u,N}$ and $\mathcal{V}^{(1+\varepsilon)u}$ such that

$$\lim_{N \rightarrow \infty} N^\kappa \left(1 - \hat{\mathbb{P}}^{u,N} \left[\mathcal{V}^{(1+\varepsilon)u} \cap \mathbb{A} \subseteq \Phi_N(\mathcal{V}^{u,N} \cap A) \right] \right) = 0. \quad (2.19)$$

Now we fix $u > 0$ and $\varepsilon > 0$ so that the distribution of $\mathcal{V}^{(1+\varepsilon)u}$ satisfies (1.8) and (1.11). Let $u' = (1 + \varepsilon)u$, let $\rho_{u'}(\cdot, \cdot)$ be the chemical distance in $\mathcal{V}^{u'}$, and let $C_{u'} < \infty$ be the constant from (1.11). Let $\mathcal{V}_\infty^{u'}$ be the unique (by (1.11)) infinite connected component of $\mathcal{V}^{u'}$. Consider the event

$$\mathcal{E}^{u',N} = \left\{ \begin{array}{l} \text{for any } x, y \in \mathcal{V}_\infty^{u'} \cap B(0, n), \rho_{u'}(x, y) \leq C_{u'}n, \text{ and} \\ \text{for any } x \in B(0, n/2), B(x, n^{1/d}) \cap \mathcal{V}_\infty^{u'} \neq \emptyset \end{array} \right\}.$$

It follows from (1.8), (1.11), and our choice of u' that for any $\kappa > 0$,

$$\lim_{N \rightarrow \infty} N^\kappa \left(1 - \hat{\mathbb{P}}^{u,N} \left[\mathcal{E}^{u',N} \right] \right) = 0. \quad (2.20)$$

Note that if the event $\mathcal{E}^{u',N} \cap \{\mathcal{V}^{u'} \cap A \subseteq \Phi_N(\mathcal{V}^{u,N} \cap A)\}$ occurs, then the set

$$\mathcal{C} = \Phi_N^{-1} \left(\mathcal{V}_\infty^{u'} \cap B(0, n) \right) \subseteq \mathcal{V}^{u,N} \cap B(0, n)$$

satisfies the conditions in (2.18) for all large enough N . Indeed, it obviously satisfies (ii) by the definition of $\mathcal{E}^{u',N}$ and the fact that $B(x, n^{1/d}) \subset B(0, n)$ for all $x \in B(0, n/2)$. It satisfies (i) since for any $x, y \in \mathcal{V}_\infty^{u'} \cap B(0, n)$, by the definition of $\mathcal{E}^{u',N}$, there exists a path of length at most $C_{u'}n$ in $\mathcal{V}^{u'}$ between x and y , which (if $n + C_{u'}n < n^2$) is contained in $A \cap \mathcal{V}^{u'}$ for large enough N . Therefore, there must be a path in $\mathcal{V}^{u,N}$ of length at most $C_{u'}n$ between any pair of vertices $x, y \in \mathcal{C}$ (just take the preimage of the corresponding path in $\mathcal{V}_\infty^{u'}$ under Φ_N).

Lemma 2.10 now follows from (2.19) and (2.20). \square

Corollary 2.12. *Let $d \geq 3$ and $u > 0$. Choose $\varepsilon = \varepsilon(u) > 0$ such that $(1 + \varepsilon)u < \bar{u}$, where \bar{u} is defined in Theorem 2.3 (see also Remark 2.11). Let $C = C(u) = C(u, \varepsilon(u)) < \infty$ be the constant from (2.18). Consider the event*

$$\mathcal{F}^{u,N} = \left\{ \begin{array}{l} \text{for any } z \in \mathbb{T}_N, \text{ there exists } \mathcal{C}_z \subseteq \mathcal{V}^{u,N} \cap B(z, n) \text{ such that} \\ \text{(i) for all } x, y \in \mathcal{C}_z, \rho_{u,N}(x, y) \leq Cn, \text{ and} \\ \text{(ii) for all } x \in B(z, n/2), B(x, n^{1/d}) \cap \mathcal{C}_z \neq \emptyset \end{array} \right\}. \quad (2.21)$$

It follows from Lemma 2.10 and Remark 2.11 that for any $d \geq 3$, $u < \bar{u}$, and $\kappa > 0$,

$$\lim_{N \rightarrow \infty} N^\kappa (1 - \mathbb{P}^{u,N} [\mathcal{F}^{u,N}]) = 0. \quad (2.22)$$

Lemma 2.10 establishes under some assumptions on u the existence of a subset \mathcal{C} of $\mathcal{V}^{u,N}$ with good chemical distance properties. Those assumptions are not enough to extend the result of Lemma 2.10 to \mathcal{C}_{max}^u , and we need to assume that $\mathcal{V}^{u,N}$ is well connected locally. More precisely, let $\mathcal{V}_r^{u,N}$ be the subset of vertices in connected components of $\mathcal{V}^{u,N}$ with diameter $\geq r$. Consider the event

$$\mathcal{G}^{u,N} = \left\{ \begin{array}{l} \text{for all } x, y \in \mathcal{V}_{n^{1/d}}^{u,N} \text{ with } |x - y| \leq 2n^{1/d} \\ x \text{ and } y \text{ are connected in } \mathcal{V}^{u,N} \cap B(x, 4n^{1/d}) \end{array} \right\}.$$

The following lemma provides sufficient conditions for the connectedness and ubiquity of $\mathcal{V}_{n^{1/d}}^{u,N}$ and states that the (chemical) diameter of $\mathcal{V}_{n^{1/d}}^{u,N}$ is comparable to N .

Lemma 2.13. *For any $d \geq 3$ and $u > 0$, there exists $C' = C'(u) < \infty$ and $N' < \infty$ such that for all $N \geq N'$,*

$$\mathcal{F}^{u,N} \cap \mathcal{G}^{u,N} \subseteq \left\{ \begin{array}{l} \text{for all } x, y \in \mathcal{V}_{n^{1/d}}^{u,N}, \rho_{u,N}(x, y) \leq C'N, \text{ and} \\ \text{for all } x \in \mathbb{T}_N, \mathcal{V}_{n^{1/d}}^{u,N} \cap B(x, n^{1/d}) \neq \emptyset \end{array} \right\}. \quad (2.23)$$

Proof of Lemma 2.13. Fix $d \geq 3$ and $u > 0$, and assume that the event $\mathcal{F}^{u,N} \cap \mathcal{G}^{u,N}$ occurs. First of all, note that for any $z \in \mathbb{T}_N$, by the definition (2.21) of \mathcal{C}_z ,

$$\mathcal{C}_z \subset \mathcal{V}_{n^{1/d}}^{u,N}. \quad (2.24)$$

Indeed, \mathcal{C}_z is a subset of a connected set in $\mathcal{V}^{u,N}$ and contains vertices at distance $\geq n/2 - 2n^{1/d} > n^{1/d}$ (the last inequality holds if N is large enough). It follows from (2.24) and the definitions of $\mathcal{F}^{u,N}$ and $\mathcal{G}^{u,N}$ that

$$\text{for all } x \in \mathcal{V}_{n^{1/d}}^{u,N}, \text{ there exists } y \in \mathcal{C}_x \text{ such that } \rho_{u,N}(x, y) \leq |B(x, 4n^{1/d})| \leq 9^d \cdot n. \quad (2.25)$$

For $x, y \in \mathbb{T}_N$, let

$$\rho_{u,N}(\mathcal{C}_x, \mathcal{C}_y) = \inf\{\rho_{u,N}(x', y') : x' \in \mathcal{C}_x, y' \in \mathcal{C}_y\} \in [0, \infty].$$

Note that

$$\text{for any } x, y \in \mathbb{T}_N \text{ with } |x - y| < n, \quad \rho_{u,N}(\mathcal{C}_x, \mathcal{C}_y) \leq 9^d \cdot n. \quad (2.26)$$

Indeed, for any $x, y \in \mathbb{T}_N$ with $|x - y| < n$, there exists $z \in B(x, n/2) \cap B(y, n/2)$, therefore there exist $x' \in \mathcal{C}_x \cap B(z, n^{1/d})$ and $y' \in \mathcal{C}_y \cap B(z, n^{1/d})$. By (2.24), the definition of $\mathcal{G}^{u,N}$, and the fact that $|x' - y'| \leq 2n^{1/d}$, $\rho_{u,N}(x', y') \leq |B(x', 4n^{1/d})| \leq 9^d \cdot n$. This is precisely (2.26). It is now immediate from (2.26) that

$$\text{for any } x, y \in \mathbb{T}_N, \quad \rho_{u,N}(\mathcal{C}_x, \mathcal{C}_y) \leq 2 \cdot (9^d + C) \cdot N, \quad (2.27)$$

where $C = C(u)$ is the constant in (2.21). Indeed, for any $x, y \in \mathbb{T}_N$, there exist $z_1, \dots, z_m \in \mathbb{T}_N$ such that $m = 2\lfloor N/n \rfloor$, $z_1 = x$, $z_m = y$, and for all $1 \leq i < m$, $|z_i - z_{i+1}| < n$. Applying (2.26) to each pair z_i, z_{i+1} and using the definition of \mathcal{C}_{z_i} , we obtain from the triangle inequality that

$$\rho_{u,N}(\mathcal{C}_x, \mathcal{C}_y) \leq \sum_{i=1}^{m-1} \rho_{u,N}(\mathcal{C}_{z_i}, \mathcal{C}_{z_{i+1}}) + (m-2) \cdot C \cdot n \leq 2 \cdot (9^d + C) \cdot N,$$

where $C = C(u)$ is the constant in (2.21).

It follows, in particular, from (2.27) that all the \mathcal{C}_z , $z \in \mathbb{T}_N$, are connected in $\mathcal{V}^{u,N}$. The event on the right-hand side of (2.23) now follows from (2.25), (2.27), and the triangle inequality, with $C' = 4 \cdot (9^d + C)$. \square

Corollary 2.14. *Let \mathcal{C}_{max}^u be a connected component in $\mathcal{V}^{u,N}$ with the largest volume. It is immediate that if $\mathcal{V}_{n^{1/d}}^{u,N}$ is connected and has diameter greater than n , then it is the connected component in $\mathcal{V}^{u,N}$ with the largest volume, i.e., $\mathcal{C}_{max}^u = \mathcal{V}_{n^{1/d}}^u$. Therefore, (2.23) implies that*

$$\mathcal{F}^{u,N} \cap \mathcal{G}^{u,N} \subseteq \left\{ \begin{array}{l} \text{for all } x, y \in \mathcal{C}_{max}^u, \rho_{u,N}(x, y) \leq C'N, \text{ and} \\ \text{for all } x \in \mathbb{T}_N, \mathcal{C}_{max}^u \cap B(x, n^{1/d}) \neq \emptyset \end{array} \right\}, \quad (2.28)$$

with C' as in (2.23).

Now we have all the ingredients to prove Theorem 2.8.

Proof of Theorem 2.8. In [30, Definition 2.4] a notion of strongly supercritical values of u was introduced, which is similar in spirit to assumption **S1** for \mathcal{V}^u , but stronger. In particular, it is not yet known that strongly supercritical u 's exist for $d \in \{3, 4\}$. We will not give this definition here, but just mention that analogously to [30, (2.20)-(2.22)] one obtains that for any strongly supercritical u and for any $\kappa > 0$,

$$\lim_{N \rightarrow \infty} N^\kappa (1 - \mathbb{P}^{u,N}[\mathcal{G}^{u,N}]) = 0. \quad (2.29)$$

In particular, we conclude from (2.22), (2.28), and (2.29) that for any $u < \bar{u}$ which is strongly supercritical, (2.17) holds. It remains to say that by [29, Theorems 3.2 and 3.3] (see [30, (1.11)]),

for all $d \geq 5$, there exists $\hat{u} > 0$ such that every $u \in (0, \hat{u})$ is strongly supercritical.

This completes the proof of Theorem 2.8 with $\tilde{u} = \min(\bar{u}, \hat{u})$. \square

3 Renormalization

In this short section we set up a general multi-scale renormalization scheme and state Theorem 3.1.

Let $l_k, r_k, L_k, k \geq 0$ be sequences of positive integers such that

$$l_k > r_k, \quad L_k = l_{k-1} \cdot L_{k-1}, \quad k \geq 1. \quad (3.1)$$

For $k \geq 0$, we introduce the renormalized lattice graph \mathbb{G}_k by

$$\mathbb{G}_k = L_k \mathbb{Z}^d = \{L_k x : x \in \mathbb{Z}^d\},$$

with edges between any pair of (l^1 -)nearest neighbor vertices of \mathbb{G}_k . For $k \geq 1$ and $x \in \mathbb{G}_k$, let

$$\Lambda_{x,k} = \mathbb{G}_{k-1} \cap (x + [0, L_k)^d). \quad \left(\text{Note that } |\Lambda_{x,k}| = (l_{k-1})^d. \right) \quad (3.2)$$

For $x \in \mathbb{G}_0$, let $\overline{G}_x = \overline{G}_{x,0} = \overline{G}_{x,0,L_0}$ be a $\sigma(\Psi_y, y \in x + [-L_0, 3L_0]^d)$ -measurable event. We call events of the form $\overline{G}_{x,0,L_0}$ *seed events*, and we denote the family of seed events $(\overline{G}_{x,0,L_0} : L_0 \geq 1, x \in \mathbb{G}_0)$ by \overline{G} . For $k \geq 1$ and $x \in \mathbb{G}_k$, we recursively define the events $\overline{G}_{x,k} = \overline{G}_{x,k,L_0}$ by

$$\overline{G}_{x,k} = \bigcup_{x_1, x_2 \in \Lambda_{x,k}; |x_1 - x_2|_\infty > r_{k-1} \cdot L_{k-1}} \overline{G}_{x_1,k-1} \cap \overline{G}_{x_2,k-1}. \quad (3.3)$$

(For simplicity, we omit the dependence of $\overline{G}_{x,k}$ on L_0 from the notation.) The reader should think of the event $\overline{G}_{x,k}$ as “unfavorable”. Informally, (3.3) states that an unfavorable event inside a box of scale L_k implies the occurrence of two unfavorable events inside sub-boxes of scale L_{k-1} that are sufficiently far apart from each other. By induction on k , the event $\overline{G}_{x,k}$ is $\sigma(\Psi_y, y \in x + [-L_0, L_k + 2L_0]^d)$ -measurable.

The following theorem helps us to give a good upper bound on the \mathbb{P}^u -probability of some large-scale unfavorable events $\overline{G}_{x,k}, x \in \mathbb{G}_k$ in the case when the events $\overline{G}_x, x \in \mathbb{G}_0$, are $\mathbb{P}^{u'}$ -unlikely for any u' in a small neighborhood of u .

Theorem 3.1. *Assume that the probability measures $\mathbb{P}^u, u \in (a, b)$, satisfy conditions **P2** and **P3**. Let $u \in (a, b)$ and $\delta \in (0, 1)$. Let*

$$\theta_{sc} = \lceil 1/\varepsilon_P \rceil, \quad l_k = l_0 \cdot 4^{k\theta_{sc}}, \quad r_k = r_0 \cdot 2^{k\theta_{sc}}, \quad k \geq 0, \quad (3.4)$$

where ε_P is defined in **P3**. Let $\overline{G}_x, x \in \mathbb{G}_0$, be all increasing events and $u' = (1 + \delta)u$, or all decreasing events and $u' = (1 - \delta)u$. If $u' \in (a, b)$ and

$$\liminf_{L_0 \rightarrow \infty} \sup_{x \in \mathbb{G}_0} \mathbb{P}^{u'} [\overline{G}_x] = 0, \quad (3.5)$$

then there exist $C = C(\delta) < \infty$ and L_0 such that for any choice of $l_0, r_0 \geq C$ and for all $k \geq 0$,

$$\sup_{x \in \mathbb{G}_k} \mathbb{P}^u [\overline{G}_{x,k}] \leq 2^{-2^k}. \quad (3.6)$$

Moreover, if the limit in (3.5) (as $L_0 \rightarrow \infty$) exists and equals to 0, then there exists $C' = C'(\delta, \overline{G}, l_0) < \infty$ such that the inequality (3.6) holds for all $l_0, r_0 \geq C', L_0 \geq C'$, and $k \geq 0$.

Remark 3.2. The choice of l_k and r_k in (3.4) is somewhat arbitrary and made to make some of the calculations more transparent, nevertheless, for the proof of Theorem 3.1, we want that

- (a) $\prod_{k=0}^{\infty} \left(1 + r_k^{-\chi_P}\right)$ is close to 1, where χ_P is defined in (1.1) (see (7.1)),
- (b) l_k grow fast enough (see (7.11), where the choice of θ_{sc} becomes clear), but
- (c) l_k do not grow too rapidly (see (7.10)).

For the proof of Theorem 1.3, we will also want that

- (d) $\prod_{k=1}^{\infty} \left(1 + \frac{r_{k-1}}{l_{k-1}}\right) < \infty$ (see (6.5)),
- (e) $l_0 > 4r_0$ (see Lemma 5.3), and
- (f) l_k do not grow too fast (see (6.2) and (6.3)).

We postpone the proof of Theorem 3.1 until Section 7. Our particular choice of the monotone seed events (and the verification of the condition (3.5) for these seed events) is the subject of Section 4.

4 Connectivity patterns in \mathcal{S}

The strategy for the proof of Theorem 1.3 is to define a \mathbb{P}^u -likely event \mathcal{H} , which implies the event on the left-hand side of (1.11). The definition of \mathcal{H} in Section 6 will be based on a certain coarse-graining of \mathcal{S} . We will partition the vertices of the coarse-grained lattice \mathbb{G}_0 into 0-good and 0-bad in such a way that

- (a) if x is a 0-good vertex of \mathbb{G}_0 then we can identify a macroscopic connected component \mathcal{C}_x of $\mathcal{S} \cap (x + [0, L_0]^d)$ with the property that components $\mathcal{C}_x, \mathcal{C}_y$ corresponding to adjacent good vertices $x, y \in \mathbb{G}_0$ are locally connected to each other in \mathcal{S} , and
- (b) the subgraph of \mathbb{G}_0 spanned by good vertices contains a dense network of “highways”, i.e., nearest neighbor paths whose length is comparable to the graph distance between their endpoints in \mathbb{G}_0 .

The crucial role in the definition of the event \mathcal{H} will be played by Theorem 3.1. Since Theorem 3.1 is only about increasing or decreasing events, we will define the event that $x \in \mathbb{G}_0$ is 0-good as an intersection of an increasing event A_x^u and a decreasing event B_x^u . The event A_x^u roughly asserts that the L_0 -neighborhood of x has a macroscopic connected component of \mathcal{S} (see Section 4.1), and the event B_x^u roughly asserts that the L_0 -neighborhood of x has at most one macroscopic connected component of \mathcal{S} (see Section 4.2).

The geometric details of the above listed plan will be carried out in Section 5. The property of 0-good vertices outlined in (a) will be precisely formulated and proved in Lemma 5.2. We will obtain the construction of the “highways” mentioned in (b) by iteratively applying the construction of Lemma 5.3.

Before we proceed with the definition and properties of the events A_x^u and B_x^u , observe that by (1.10) and Definition 1.5, $\mathbb{P}^u[0 \in \mathcal{S}_r] \geq \eta(u)$ and $\mathbb{P}^u[0 \in \mathcal{S}_r] \rightarrow \eta(u)$ as $r \rightarrow \infty$. Therefore, if the measures \mathbb{P}^u , $u \in (a, b)$, satisfy the condition **P1**, then by an appropriate ergodic theorem (see, e.g., [10, Theorem VIII.6.9]), we get for any $u \in (a, b)$,

$$\lim_{L \rightarrow \infty} \frac{1}{L^d} \sum_{x \in [0, L]^d} \mathbb{1}(x \in \mathcal{S}_L) \stackrel{\mathbb{P}^u\text{-a.s.}}{=} \lim_{L \rightarrow \infty} \frac{1}{L^d} \sum_{x \in [0, L]^d} \mathbb{1}(x \in \mathcal{S}_{\infty}) \stackrel{\mathbb{P}^u\text{-a.s.}}{=} \eta(u), \quad (4.1)$$

where the $\{0, 1\}$ -valued random variable $\mathbb{1}(A)$ denotes the indicator of the event $A \in \mathcal{F}$.

4.1 The increasing events A_x^u

Definition 4.1. For $x \in \mathbb{G}_0$ and $u > 0$, let $A_x^u \in \mathcal{F}$ be the event that

- (a) for each $e \in \{0, 1\}^d$, the set $\mathcal{S}_{L_0} \cap (x + eL_0 + [0, L_0)^d)$ contains a connected component with at least $\frac{3}{4}\eta(u)L_0^d$ vertices,
- (b) all of these 2^d components are connected in $\mathcal{S} \cap (x + [0, 2L_0)^d)$.

Note that for each $y' \in \mathbb{Z}^d$, the event $\{y' \in \mathcal{S}_{L_0}\}$ is increasing and measurable with respect to $\sigma(\Psi_y : y \in B(y', L_0))$, thus A_x^u is increasing and measurable with respect to $\sigma(\Psi_y : y \in x + [-L_0, 3L_0)^d)$.

For $u > 0$ and $x \in \mathbb{G}_0$, let

$$\overline{A}_{x,0}^u = (A_x^u)^c,$$

and for $u > 0$, $k \geq 1$, and $x \in \mathbb{G}_k$ let

$$\overline{A}_{x,k}^u = \bigcup_{x_1, x_2 \in \Lambda_{x,k}; |x_1 - x_2|_\infty > r_{k-1} \cdot L_{k-1}} \overline{A}_{x_1, k-1}^u \cap \overline{A}_{x_2, k-1}^u. \quad (4.2)$$

The events $\overline{A}_{x,k}^u$ are decreasing and measurable with respect to $\sigma(\Psi_y, y \in x + [-L_0, L_k + 2L_0)^d)$.

Lemma 4.2. Assume that the measures \mathbb{P}^u , $u \in (a, b)$, satisfy conditions **P1** – **P3** and **S1** – **S2**. Let l_k , r_k , and L_k be defined as in (3.1) and (3.4). Then for each $u \in (a, b)$, there exist $C = C(u) < \infty$ and $C' = C'(u, l_0) < \infty$ such that for all $l_0, r_0 \geq C$, $L_0 \geq C'$, and $k \geq 0$,

$$\mathbb{P}^u [\overline{A}_{0,k}^u] \leq 2^{-2^k}.$$

Proof of Lemma 4.2. It suffices to show that for any $u \in (a, b)$ there exists $\delta = \delta(u) > 0$ such that $(1 - \delta)u > a$ and

$$\mathbb{P}^{(1-\delta)u} [A_0^u] \rightarrow 1, \quad \text{as } L_0 \rightarrow \infty. \quad (4.3)$$

Indeed, once (4.3) is shown, the result will follow from Theorem 3.1 applied to the decreasing seed events $\overline{G}_x = \overline{A}_{x,0}^u$, $x \in \mathbb{G}_0$.

Let $u \in (a, b)$. By the condition **S2**, we can choose $\varepsilon > 0$ and $\delta = \delta(u) > 0$ so that if we define $u' = (1 - \delta)u$, then we have

$$u' > a \quad \text{and} \quad (1 - 4\varepsilon)^d \eta(u') > \frac{3}{4}\eta(u). \quad (4.4)$$

With such a choice of ε and δ , for $L_0 \geq 1$, we obtain

$$\eta(u')(L_0 - 4\lfloor \varepsilon L_0 \rfloor)^d > \frac{3}{4}\eta(u)L_0^d.$$

We consider the boxes

$$Q_e = eL_0 + [2\lfloor \varepsilon L_0 \rfloor, L_0 - 2\lfloor \varepsilon L_0 \rfloor)^d, \quad e \in \{0, 1\}^d.$$

The volume of Q_e is $|Q_e| = (L_0 - 4\lfloor \varepsilon L_0 \rfloor)^d$. Using (4.1) and (4.4), we get that with $\mathbb{P}^{u'}$ -probability tending to 1 as $L_0 \rightarrow \infty$, each of the boxes Q_e , $e \in \{0, 1\}^d$ contains at least $\frac{3}{4}\eta(u)L_0^d$ vertices of \mathcal{S}_{L_0} .

By (1.9), we obtain that with $\mathbb{P}^{u'}$ -probability tending to 1 as $L_0 \rightarrow \infty$, for all $e \in \{0, 1\}^d$, $\mathcal{S}_{L_0} \cap Q_e$ is a subset of a connected component $\tilde{\mathcal{C}}_e$ of

$$\mathcal{S}_{L_0} \cap \left(eL_0 + [\lfloor \varepsilon L_0 \rfloor, L_0 - \lfloor \varepsilon L_0 \rfloor)^d \right).$$

By what is proved earlier, with $\mathbb{P}^{u'}$ -probability tending to 1 as $L_0 \rightarrow \infty$, for all $e \in \{0, 1\}^d$, we have $|\tilde{\mathcal{C}}_e| \geq \frac{3}{4}\eta(u)L_0^d$.

Moreover, again by (1.9), with $\mathbb{P}^{u'}$ -probability tending to 1 as $L_0 \rightarrow \infty$, all $\tilde{\mathcal{C}}_e$, $e \in \{0, 1\}^d$, are in the same connected component of $\mathcal{S}_{L_0} \cap [0, 2L_0)^d$.

By Definition 4.1, this implies (4.3) and completes the proof of Lemma 4.2. \square

4.2 The decreasing events B_x^u

Definition 4.3. For $x \in \mathbb{G}_0$ and $u > 0$, let $B_x^u \in \mathcal{F}$ be the event that for all $e \in \{0, 1\}^d$,

$$\left| \mathcal{S}_{L_0} \cap (x + eL_0 + [0, L_0)^d) \right| \leq \frac{5}{4} \eta(u) L_0^d.$$

Note that the event B_x^u is decreasing and measurable with respect to $\sigma(\Psi_y : y \in x + [-L_0, 3L_0]^d)$.

For $u > 0$ and $x \in \mathbb{G}_0$, let

$$\overline{B}_{x,0}^u = (B_x^u)^c,$$

and for $u > 0$, $k \geq 1$, and $x \in \mathbb{G}_k$ let

$$\overline{B}_{x,k}^u = \bigcup_{x_1, x_2 \in \Lambda_{x,k}; |x_1 - x_2|_\infty > r_{k-1} \cdot L_{k-1}} \overline{B}_{x_1, k-1}^u \cap \overline{B}_{x_2, k-1}^u. \quad (4.5)$$

The events $\overline{B}_{x,k}^u$ are increasing and measurable with respect to $\sigma(\Psi_y, y \in x + [-L_0, L_k + 2L_0]^d)$.

Lemma 4.4. Assume that the measures \mathbb{P}^u , $u \in (a, b)$, satisfy conditions **P1** – **P3** and **S2**. Let l_k , r_k , and L_k be defined as in (3.1) and (3.4). For each $u \in (a, b)$, there exist $C = C(u) < \infty$ and $C' = C'(u, l_0) < \infty$ such that for all $l_0, r_0 \geq C$, $L_0 \geq C'$, and $k \geq 0$,

$$\mathbb{P}^u [\overline{B}_{0,k}^u] \leq 2^{-2^k}.$$

Proof of Lemma 4.4. It suffices to show that for any $u \in (a, b)$ there exists $\delta = \delta(u) > 0$ such that $(1 + \delta)u < b$ and

$$\mathbb{P}^{(1+\delta)u} [B_0^u] \rightarrow 1, \quad \text{as } L_0 \rightarrow \infty. \quad (4.6)$$

Indeed, once (4.6) is shown, the result will follow from Theorem 3.1 applied to the increasing seed events $\overline{G}_x = \overline{B}_{x,0}^u$, $x \in \mathbb{G}_0$.

Let $u \in (a, b)$. By the condition **S2**, we can choose $\delta = \delta(u) > 0$ so that with $u' = (1 + \delta)u$, we have

$$u' < b \quad \text{and} \quad \eta(u') < \frac{5}{4} \eta(u). \quad (4.7)$$

Therefore, (4.1) and (4.7) imply that, with $\mathbb{P}^{u'}$ probability tending to 1 as $L_0 \rightarrow \infty$, for all $e \in \{0, 1\}^d$,

$$\left| \mathcal{S}_{L_0} \cap (eL_0 + [0, L_0)^d) \right| < \frac{5}{4} \eta(u) L_0^d.$$

By Definition 4.3, this is precisely (4.6). The proof of Lemma 4.4 is complete. \square

5 Good and bad vertices

The aim of this section is to provide the geometric details of one iteration of the recursive construction of short paths in \mathcal{S} . We follow the strategy outlined in the introduction of Section 4. Throughout this section, l_k , r_k , and L_k , $k \geq 1$, are positive integers satisfying (3.1).

Definition 5.1. Let $u \in (a, b)$. For $k \geq 0$, we say that $x \in \mathbb{G}_k$ is *k-bad* if the event

$$\overline{A}_{x,k}^u \cup \overline{B}_{x,k}^u$$

occurs. Otherwise, we say that x is *k-good*. Note that the event $\{x \text{ is } k\text{-good}\}$ is measurable with respect to $\sigma(\Psi_y : y \in x + [-L_0, L_k + 2L_0]^d)$, see below (4.2) and (4.5).

In the remainder of this section we prove useful geometric properties of k -good vertices. All the results of this section are purely deterministic. We fix the parameter $u \in (a, b)$.

Lemma 5.2. *Let x and y be nearest neighbors in \mathbb{G}_0 and assume that they are both 0-good.*

- (a) *Each of the graphs $\mathcal{S}_{L_0} \cap (z + [0, L_0]^d)$, with $z \in \{x, y\}$, contains the unique connected component \mathcal{C}_z with at least $\frac{3}{4}\eta(u)L_0^d$ vertices, and*
- (b) *\mathcal{C}_x and \mathcal{C}_y are connected in the graph $\mathcal{S} \cap ((x + [0, 2L_0]^d) \cup (y + [0, 2L_0]^d))$.*

Proof of Lemma 5.2. Let x and y be nearest neighbors in \mathbb{G}_0 and assume they are both 0-good, i.e., the events A_x^u , B_x^u , A_y^u , and B_y^u occur.

By Definition 4.1, since A_x^u and A_y^u occur, the graphs $\mathcal{S}_{L_0} \cap (x + [0, L_0]^d)$ and $\mathcal{S}_{L_0} \cap (y + [0, L_0]^d)$ contain connected components \mathcal{C}_x and \mathcal{C}_y with at least $\frac{3}{4}\eta(u)L_0^d$ vertices, which are connected in $\mathcal{S} \cap ((x + [0, 2L_0]^d) \cup (y + [0, 2L_0]^d))$.

It remains to show that \mathcal{C}_x and \mathcal{C}_y are unique. By Definition 4.3, since B_x^u and B_y^u occur, each of the sets $\mathcal{S}_{L_0} \cap (x + [0, L_0]^d)$ and $\mathcal{S}_{L_0} \cap (y + [0, L_0]^d)$ contain at most $\frac{5}{4}\eta(u)L_0^d$ vertices. Since $2 \cdot \frac{3}{4} > \frac{5}{4}$, there can be at most one connected component of size at least $\frac{3}{4}\eta(u)L_0^d$ in each of the graphs $\mathcal{S}_{L_0} \cap (x + [0, L_0]^d)$ and $\mathcal{S}_{L_0} \cap (y + [0, L_0]^d)$. This implies the uniqueness of \mathcal{C}_x and \mathcal{C}_y and finishes the proof of Lemma 5.2. \square

Lemma 5.3. *Let $k \geq 1$. Let $\pi = (x_0, \dots, x_m)$ be a nearest neighbor path in \mathbb{G}_k of k -good vertices. If $l_0 > 4r_0$, then there exists a nearest neighbor path $\pi' = (x'_0, \dots, x'_n)$ in \mathbb{G}_{k-1} of $(k-1)$ -good vertices, such that*

- (a) $x'_0 \in \Lambda_{x_0, k}$, $x'_n \in \Lambda_{x_m, k}$ (recall (3.2)), and
- (b) $n \leq (1 + 8 \cdot \frac{r_{k-1}}{l_{k-1}}) \cdot l_{k-1}m$.

Proof of Lemma 5.3. Fix $k \geq 1$ and a nearest neighbor path $\pi = (x_0, \dots, x_m)$ in \mathbb{G}_k of k -good vertices. For each $0 \leq i \leq m$, since x_i is k -good (see Definition 5.1, (4.2), and (4.5)), there exist $a_i, b_i \in \Lambda_{x_i, k}$ such that all the vertices contained in

$$\Lambda_i^* = \Lambda_{x_i, k} \setminus \left((a_i + [0, r_{k-1}L_{k-1}]^d) \cup (b_i + [0, r_{k-1}L_{k-1}]^d) \right) \quad (5.1)$$

are $(k-1)$ -good. Since $l_{k-1} > r_{k-1}$, each of the sets Λ_i^* , $0 \leq i \leq m$ is non-empty. In particular, this settles the case $m = 0$, and from now on we may and will assume that $m \geq 1$.

Our strategy to construct the path π' is as follows. We show that for any $0 \leq i \leq m-1$, there exists a pair of vertices $y_i^* \in \Lambda_i^*$ and $z_{i+1}^* \in \Lambda_{i+1}^*$ such that

- (a) for $0 \leq i \leq m-1$, y_i^* is connected to z_{i+1}^* by a nearest neighbor path in $\Lambda_i^* \cup \Lambda_{i+1}^*$ of length l_{k-1} ,
- (b) for $1 \leq i \leq m-1$, z_i^* is connected to y_i^* by a nearest neighbor path in Λ_i^* of length at most $8r_{k-1}$.

We then define π' as the nearest neighbor path from $y_0^* \in \Lambda_1^*$ to $z_m^* \in \Lambda_m^*$ obtained by concatenating all the paths constructed in (a) and (b).

Let $0 \leq i \leq m-1$. Let α_i be the smallest number in $\{1, \dots, d\}$ such that e_{α_i} is orthogonal to $(x_{i+1} - x_i)$. By (5.1), since $l_{k-1} > 4r_{k-1}$, there exists an integer $0 \leq j_i \leq 4r_{k-1}$ such that the hyperplane H_i^* orthogonal to e_{α_i} and passing through $x_i + L_{k-1}j_i e_{\alpha_i} \in \Lambda_{x_i, k}$ (and $x_{i+1} + L_{k-1}j_i e_{\alpha_i} \in \Lambda_{x_{i+1}, k}$) satisfies

$$H_i^* \cap (\Lambda_{x_i, k} \cup \Lambda_{x_{i+1}, k}) \subset (\Lambda_i^* \cup \Lambda_{i+1}^*).$$

We define

$$y_i^* := x_i + L_{k-1}j_i e_{\alpha_i} \quad \text{and} \quad z_{i+1}^* := x_{i+1} + L_{k-1}j_i e_{\alpha_i}.$$

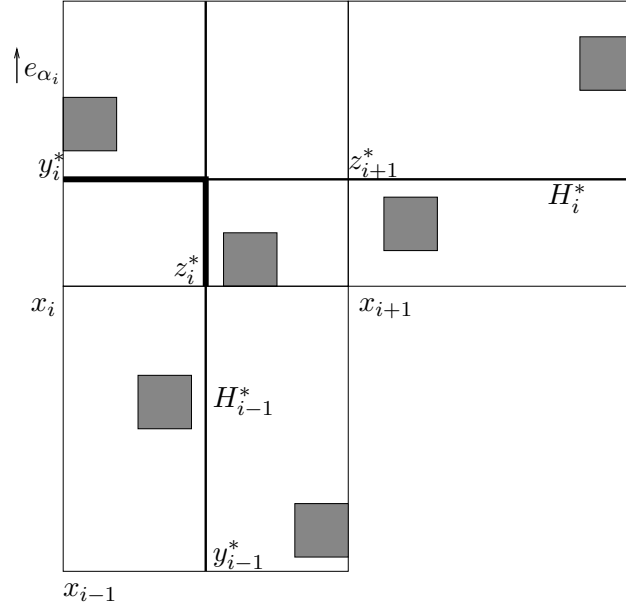


Figure 1: The case $d = 2$ and $\alpha_i \neq \alpha_{i-1}$.

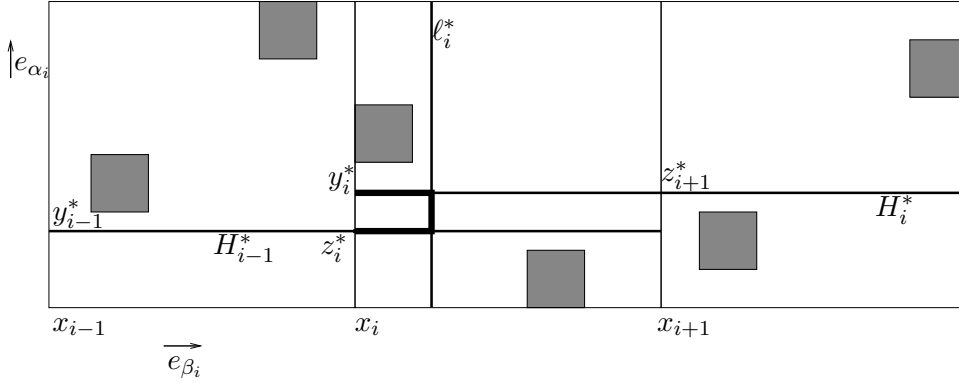


Figure 2: The case $d = 2$ and $\alpha_i = \alpha_{i-1}$.

Note that the line segment in \mathbb{G}_{k-1} from y_i^* to z_{i+1}^* has length l_{k-1} and is contained in $\Lambda_i^* \cup \Lambda_{i+1}^*$. Therefore, it remains to show that y_i^* and z_i^* satisfy the requirement (b) above.

Let $1 \leq i \leq m-1$. We consider separately the cases $\alpha_i \neq \alpha_{i-1}$ and $\alpha_i = \alpha_{i-1}$ (see Figures 1 and 2, respectively).

If $\alpha_i \neq \alpha_{i-1}$, then

$$y_i^* + L_{k-1}j_{i-1}e_{\alpha_{i-1}} = z_i^* + L_{k-1}j_ie_{\alpha_i} \in H_i^* \cap H_{i-1}^*.$$

Note that the line segment from y_i^* to $y_i^* + L_{k-1}j_{i-1}e_{\alpha_{i-1}}$ in \mathbb{G}_{k-1} is contained in $\Lambda_{x_i,k} \cap H_i^* \subset \Lambda_i^*$, and the line segment from z_i^* to $z_i^* + L_{k-1}j_ie_{\alpha_i}$ in \mathbb{G}_{k-1} is contained in $\Lambda_{x_i,k} \cap H_{i-1}^* \subset \Lambda_i^*$. We conclude that y_i^* and z_i^* are connected in Λ_i^* by a nearest neighbor path in \mathbb{G}_{k-1} of length $j_{i-1} + j_i \leq 8r_{k-1}$.

If $\alpha_i = \alpha_{i-1}$, let $\beta_i \in \{1, \dots, d\}$ be such that the unit vector e_{β_i} is parallel to $(x_i - x_{i-1})$. By the definitions of α_i and β_i , $\alpha_i \neq \beta_i$. By (5.1), since $l_{k-1} > 2r_{k-1}$, there exists an integer $0 \leq k_i \leq 2r_{k-1}$ such

that the line ℓ_i^* parallel to e_{α_i} and passing through $y_i^* + L_{k-1}k_i e_{\beta_i}$ (and $z_i^* + L_{k-1}k_i e_{\beta_i}$) satisfies

$$\ell_i^* \cap \Lambda_{x_i, k} \subset \Lambda_i^*.$$

Note that the line segment from y_i^* to $y_i^* + L_{k-1}k_i e_{\beta_i}$ in \mathbb{G}_{k-1} is contained in $\Lambda_{x_i, k} \cap H_i^* \subset \Lambda_i^*$, the line segment from z_i^* to $z_i^* + L_{k-1}k_i e_{\beta_i}$ in \mathbb{G}_{k-1} is contained in $\Lambda_{x_i, k} \cap H_{i-1}^* \subset \Lambda_i^*$, and the line segment from $y_i^* + L_{k-1}k_i e_{\beta_i}$ to $z_i^* + L_{k-1}k_i e_{\beta_i}$ in \mathbb{G}_{k-1} is contained in $\Lambda_{x_i, k} \cap \ell_i^* \subset \Lambda_i^*$. We conclude that y_i^* and z_i^* are connected in Λ_i^* by a nearest neighbor path in \mathbb{G}_{k-1} of length $2k_i + |j_{i-1} - j_i| \leq 8r_{k-1}$.

We have shown the existence of vertices y_i^* and z_i^* satisfying (a) and (b) above. The proof of Lemma 5.3 is complete. \square

6 Proof of Theorem 1.3

In our proof of Theorem 1.3 we will follow the strategy sketched at the beginning of Section 4. We will define an event \mathcal{H} using the definition of k -good vertices from Section 5. Using the seed estimates of Section 4, we will show that \mathcal{H} occurs with high \mathbb{P}^u -probability. We will then apply the geometric results of Section 5 to show that \mathcal{H} implies the event on the left-hand side of (1.11). This part of the proof is purely deterministic.

Let $u \in (a, b)$. It suffices to show that for some $c = c(u) > 0$ and $C = C(u) < \infty$, the inequality (1.11) holds for all $R \geq C'$ with large enough $C' = C'(u)$. Increasing C possibly even further then yields the validity of (1.11) for all $R \geq 1$.

We take $l_0 > 4r_0$ and L_0 so that the implications of Lemmas 4.2 and 4.4 hold (with l_k and r_k as in (3.4) and L_k as in (3.1)). Let s be the largest integer such that $L_s \leq R^{1/d}$, i.e.,

$$s = \max\{s' : L_{s'} \leq R^{1/d}\}. \quad (6.1)$$

We assume that $R \geq L_0^d$, so that s is well-defined. Let \mathcal{H} be the event that

- (a) each $z \in \mathbb{G}_s \cap B(0, 2R)$ is s -good, and
- (b) for any $z \in \mathbb{G}_s \cap B(0, 2R)$ and $x, y \in \mathcal{S}_{L_s} \cap (z + [0, 2L_s]^d)$, x is connected to y by a path in \mathcal{S} of length at most $(4L_s)^d$.

The event in part (b) of the definition of \mathcal{H} is implied by the event that

- (b') for any $z \in \mathbb{G}_s \cap B(0, 2R)$ and $x, y \in \mathcal{S}_{L_s} \cap (z + [0, 2L_s]^d)$, x is connected to y by a simple path in $\mathcal{S} \cap (z + [-L_s, 3L_s]^d)$.

By Definition 5.1, (1.9), and Lemmas 4.2 and 4.4, for any $u \in (a, b)$ and some constants $c = c(u) > 0$ and $C = C(u) < \infty$,

$$\mathbb{P}^u[\mathcal{H}^c] \leq |\mathbb{G}_s \cap B(0, 2R)| \cdot 2 \cdot 2^{-2^s} + |\mathbb{G}_s \cap B(0, 2R)| \cdot C \cdot e^{-cfs(u, L_s)}.$$

By (3.1), (3.4), and (6.1), for all $R \geq L_0^d$,

$$R^{1/d} \leq L_{s+1} = l_s \cdot L_s \leq l_0 \cdot 4 \cdot (L_s)^{1+2\theta_{sc}},$$

which implies that

$$L_s \geq \frac{1}{4l_0} R^{1/d(1+2\theta_{sc})}. \quad (6.2)$$

Using (1.6), (3.1), (3.4), and (6.2), we deduce that there exist $c' = c'(u) > 0$ and $C' = C'(u) < \infty$ such that for all $R \geq C'(u)$,

$$2^s \geq (\log R)^{1+\Delta_s} \quad \text{and} \quad f_s(u, L_s) \geq c'(\log R)^{1+\Delta_s}. \quad (6.3)$$

Note that we also have $|\mathbb{G}_s \cap B(0, 2R)| \leq (4R+1)^d$. Therefore, for any $u \in (a, b)$, there exist $c = c(u) > 0$ and $C = C(u) < \infty$ such that for all $R \geq C(u)$,

$$\mathbb{P}^u[\mathcal{H}] \geq 1 - Ce^{-c(\log R)^{1+\Delta_s}}. \quad (6.4)$$

We will now show that the occurrence of \mathcal{H} implies the event on the left-hand side of (1.11). Then (1.11) will immediately follow from (6.4).

Assume that \mathcal{H} occurs. Take $x, y \in \mathcal{S}_R \cap B(0, R)$. We will show that $\rho_s(x, y) \leq CR$ for some $C = C(u) < \infty$.

If there exists $z \in \mathbb{G}_s \cap B(0, 2R)$ such that $x, y \in (z + [0, 2L_s]^d)$, then the result follows from part (b) of the definition of \mathcal{H} and (6.1).

Therefore, we assume that there is no such z . There exist unique vertices $x', y' \in \mathbb{G}_s \cap B(0, 2R)$ such that

$$x \in \mathcal{S}_R \cap (x' + [0, L_s]^d), \quad y \in \mathcal{S}_R \cap (y' + [0, L_s]^d).$$

Moreover, $|x' - y'|_\infty \geq 2L_s$. By part (a) of the definition of \mathcal{H} , there exists a nearest neighbor path of s -good vertices in $\mathbb{G}_s \cap B(0, 2R)$ connecting x' and y' of length

$$N' \leq d[(4R+1)/L_s].$$

Applying Lemma 5.3 repeatedly for $k = s, s-1, \dots, 1$, we deduce that there exists 0-good vertices $x'' \in \mathbb{G}_0 \cap (x' + [0, L_s]^d)$ and $y'' \in \mathbb{G}_0 \cap (y' + [0, L_s]^d)$ which are connected by a nearest neighbor path of 0-good vertices in \mathbb{G}_0 of length

$$N'' \leq \prod_{k=1}^s \left(1 + 8 \cdot \frac{r_{k-1}}{l_{k-1}}\right) \cdot \frac{L_s}{L_0} \cdot N'. \quad (6.5)$$

By Lemma 5.2, there exist

$$x^* \in \mathcal{S} \cap (x'' + [0, L_0]^d), \quad y^* \in \mathcal{S} \cap (y'' + [0, L_0]^d)$$

such that x^* is connected to y^* by a path in \mathcal{S} of length

$$N^* \leq (2L_0)^d \cdot N''.$$

Note that

$$x^* \in \mathcal{S} \cap (x' + [0, L_s]^d), \quad y^* \in \mathcal{S} \cap (y' + [0, L_s]^d).$$

Since $|x' - y'|_\infty \geq 2L_s$, we have $|x^* - y^*|_\infty > L_s$. In particular,

$$x^*, y^* \in \mathcal{S}_{L_s}.$$

Since $x, y \in \mathcal{S}_R \cap B(0, R) \subseteq \mathcal{S}_{L_s} \cap B(0, R)$, $x, x^* \in (x' + [0, L_s]^d)$, and $y, y^* \in (y' + [0, L_s]^d)$, we obtain by part (b) of the definition of \mathcal{H} that $\rho_s(x, x^*) \leq (4L_s)^d$ and $\rho_s(y, y^*) \leq (4L_s)^d$. As a result, by the triangle inequality, we have

$$\rho_s(x, y) \leq N^* + 2(4L_s)^d.$$

It remains to observe that there exists $C = C(u) < \infty$ such that $N^* \leq \frac{1}{2}CR$ and $2(4L_s)^d \leq \frac{1}{2}CR$. Indeed, $N^* \leq \frac{1}{2}CR$ follows from (3.4) and the definitions of N', N'' and N^* , and $2(4L_s)^d \leq \frac{1}{2}CR$ follows from (6.1).

We have thus shown that if the event \mathcal{H} occurs, then any $x, y \in \mathcal{S}_S \cap B(0, R)$ are connected in \mathcal{S} and $\rho_s(x, y) \leq CR$. Theorem 1.3 now follows from (6.4). \square

7 Proof of Theorem 3.1

Assume that the family of probability measures \mathbb{P}^u , $u \in (a, b)$, satisfies **P2** and **P3**. The proof of Theorem 3.1 goes by induction on k . The key idea is that for a certain sequence $(u_k)_{k=1}^\infty$ of parameters the expression $\sup_{x \in \mathbb{G}_{k+1}} \mathbb{P}^{u_{k+1}} [\overline{G}_{x,k+1}]$ can be upper bounded by something which is not much bigger than $\sup_{x \in \mathbb{G}_k} \mathbb{P}^{u_k} [\overline{G}_{x,k}]^2$, see (7.8). Throughout this section we fix the parameters χ_P and ε_P appearing in (1.1) and (1.4), respectively.

We only give the proof in the case of increasing events \overline{G}_x . The case when \overline{G}_x are decreasing can be treated similarly, and we omit it.

Fix $u \in (a, b)$. Let $\delta \in (0, 1)$ be such that $u' = (1 + \delta)u \in (a, b)$. Let the sequences l_k , r_k , and L_k be defined as in (3.4) and (3.1). Take $r_0 \geq R_P$ (the constant from the assumption **P3**) large enough so that

$$\prod_{k=0}^{\infty} (1 + r_k^{-\chi_P}) \leq 1 + \delta. \quad (7.1)$$

Let \overline{G}_x , $x \in \mathbb{Z}^d$, be increasing events such that $\overline{G}_x \in \sigma(\Psi_y, y \in x + [-L_0, 3L_0]^d)$. By (3.3), for all $k \geq 0$ and $x \in \mathbb{G}_k$,

$$\overline{G}_{x,k} \text{ is increasing, and } \overline{G}_{x,k} \in \sigma(\Psi_y, y \in x + [-L_0, L_k + 2L_0]^d). \quad (7.2)$$

Let

$$u_0 = u' \quad \text{and} \quad u_k = (1 + r_k^{-\chi_P}) \cdot u_{k+1}, \quad k \geq 0. \quad (7.3)$$

By (7.1), we have

$$a < u \leq u_k < b, \quad \text{for all } k \geq 0. \quad (7.4)$$

For $k \geq 0$, let

$$\kappa_k = 1 + l_0 \cdot \sum_{i=k+1}^{\infty} i^{-2} \quad (= \kappa_{k+1} + l_0 \cdot (k+1)^{-2}). \quad (7.5)$$

By monotonicity of the events \overline{G}_x , (7.4), condition **P2**, and the fact that $\kappa_k \geq 1$, for all k , in order to prove (3.6), it suffices to show that

$$\sup_{x \in \mathbb{G}_k} \mathbb{P}^{u_k} [\overline{G}_{x,k}] \leq 2^{-\kappa_k 2^k}. \quad (7.6)$$

We will prove (7.6) by induction on k .

Induction start: By (3.5) and (7.5), for each l_0 , there exist L_0 such that

$$\sup_{x \in \mathbb{G}_0} \mathbb{P}^{u_0} [\overline{G}_{x,0}] \leq 2^{-\kappa_0}, \quad (7.7)$$

hence (7.6) holds for $k = 0$ and such $L_0 = L_0(l_0, u_0, \overline{G})$. (Later l_0 will be chosen large enough, independently of L_0 .)

Induction step: We assume that (7.6) holds for $k \geq 0$, and now prove that it also holds for $k + 1$. For each $x \in \mathbb{G}_k$, we have

$$\begin{aligned} \mathbb{P}^{u_{k+1}} [\overline{G}_{x,k+1}] &\stackrel{(3.3)}{\leq} \sum_{x_1, x_2 \in \Lambda_{x,k+1} : |x_1 - x_2|_\infty > r_k \cdot L_k} \mathbb{P}^{u_{k+1}} [\overline{G}_{x_1,k} \cap \overline{G}_{x_2,k}] \\ &\stackrel{(1.3), (7.2), (7.3)}{\leq} |\Lambda_{x,k+1}|^2 \left(\sup_{x \in \mathbb{G}_k} \mathbb{P}^{u_k} [\overline{G}_{x,k}]^2 + e^{-f_P(L_k)} \right) \end{aligned} \quad (7.8)$$

$$\stackrel{(3.2)}{\leq} l_k^{2d} \left(2^{-\kappa_k 2^{k+1}} + e^{-f_P(L_k)} \right). \quad (7.9)$$

Here, in (7.8) we applied (1.3) from the condition **P3** with $R = r_k$, $L = L_k$, $u = u_k$, and $\hat{u} = u_{k+1}$ (the requirements of condition **P3** are satisfied by the choice of $r_0 \geq R_P$, (7.2), (7.3), and (7.4)), and (7.9) follows from the induction assumption.

By the definitions of l_k in (3.4) and κ_k in (7.5), there exists $C < \infty$ such that for all $k \geq 0$ and $l_0 \geq C$,

$$\log(l_k^{2d}) - \kappa_k 2^{k+1} \leq -\kappa_{k+1} 2^{k+1} - 1, \quad (7.10)$$

and by the definitions of θ_{sc} , l_k in (3.4), L_k in (3.1), κ_k in (7.5), ε_P in (1.4), and the fact that $\theta_{sc} \cdot \varepsilon_P \geq 1$, there exists $C' = C'(l_0) < \infty$ such that for all $k \geq 0$, and $L_0 \geq C'$,

$$\log(l_k^{2d}) - f_P(L_k) \leq -\kappa_{k+1} 2^{k+1} - 1. \quad (7.11)$$

Plugging (7.10) and (7.11) into (7.9), we obtain that (7.6) holds for $k+1$. We finish the proof by specifying the order in which we choose large constants r_0 , l_0 and L_0 . We first choose $r_0 \geq R_P$ (the constant from the condition **P3**) so that (7.1) holds. We then choose l_0 so that (7.10) holds. In other words, there exists a finite constant $C = C(\delta) \geq R_P$ such that for all $l_0, r_0 \geq C$, (7.1) and (7.10) hold. Finally, we choose L_0 to satisfy (7.7) and (7.11). In particular, if the limit in (3.5) exists, then there exists $C' = C'(\delta, \overline{G}, l_0) < \infty$ such that for all $L_0 \geq C'$, (7.7) and (7.11) hold. We have thus finished the proof of Theorem 3.1 in the case of increasing events \overline{G}_x . \square

8 Proof of Theorem 1.5

We begin by briefly outlining our proof of Theorem 1.5, which adapts the standard recipe for shape theorems (see, e.g. [1]) to our setting.

In Section 8.1 we define the pseudometric $\tilde{\rho}_{S_\infty}(\cdot, \cdot)$ on \mathbb{Z}^d satisfying $\tilde{\rho}_{S_\infty}(x, y) = \rho_S(x, y)$ for $x, y \in S_\infty$. This trick will allow us to work with \mathbb{P}^u rather than the more cumbersome $\mathbb{P}^u[\cdot | 0 \in S_\infty]$. After stating some properties of $\tilde{\rho}_{S_\infty}(\cdot, \cdot)$, we rephrase Theorem 1.5 in terms of $\tilde{\rho}_{S_\infty}(\cdot, \cdot)$ as Proposition 8.2.

In Section 8.2 we prove Proposition 8.2. We first use Kingman's subadditive ergodic theorem [14] to deduce in Lemma 8.3 that the large-scale behaviour of $\tilde{\rho}_{S_\infty}(0, \cdot)$ along half-lines can be \mathbb{P}^u -a.s. approximated by some norm $p_u(\cdot)$ on \mathbb{R}^d . The convex set D_u will turn out to be the unit ball with respect to this norm. The proof of Proposition 8.2 combines the result of Lemma 8.3 with a covering argument in which Theorem 1.3 serves as a preliminary bound.

8.1 A suitable pseudometric on \mathbb{Z}^d

In this section we define the pseudometric $\tilde{\rho}_{S_\infty}(\cdot, \cdot)$ on \mathbb{Z}^d satisfying $\tilde{\rho}_{S_\infty}(x, y) = \rho_S(x, y)$ for $x, y \in S_\infty$. For that we assume that

$$S_\infty \neq \emptyset. \quad (8.1)$$

We first choose a bijection $\ell : \mathbb{Z}^d \rightarrow \mathbb{N}$ which satisfies

$$\forall x, y \in \mathbb{Z}^d : |x|_\infty < |y|_\infty \implies \ell(x) < \ell(y).$$

Thus ℓ gives us a labelling of \mathbb{Z}^d such that vertices with smaller norm have smaller labels. Given $x \in \mathbb{Z}^d$ and $\emptyset \neq V \subseteq \mathbb{Z}^d$ we define $\Phi(x, V) \in V$ to be one of the l^∞ -closest vertices of V to x by

$$\Phi(x, V) = x^*, \quad \text{where} \quad \ell(x^* - x) = \min\{\ell(y - x) : y \in V\}.$$

Note that

$$\Phi(x, V) = x \quad \text{for all} \quad x \in V, \quad (8.2)$$

and

$$\Phi(x + y, V + y) = \Phi(x, V) + y \quad \text{for all } x, y \in \mathbb{Z}^d. \quad (8.3)$$

We define $\tilde{\rho}_{\mathcal{S}_\infty} : \mathbb{Z}^d \times \mathbb{Z}^d \rightarrow \mathbb{Z}_+ \cup \{\infty\}$ by the formula

$$\tilde{\rho}_{\mathcal{S}_\infty}(x, y) = \rho_{\mathcal{S}}(\Phi(x, \mathcal{S}_\infty), \Phi(y, \mathcal{S}_\infty)), \quad x, y \in \mathbb{Z}^d. \quad (8.4)$$

By (8.1), the function $\tilde{\rho}_{\mathcal{S}_\infty}$ is well-defined. In fact, $\tilde{\rho}_{\mathcal{S}_\infty}(\cdot, \cdot)$ is a pseudometric on \mathbb{Z}^d . By (8.2),

$$\forall x, y \in \mathcal{S}_\infty : \tilde{\rho}_{\mathcal{S}_\infty}(x, y) = \rho_{\mathcal{S}}(x, y) \geq |x - y|_1. \quad (8.5)$$

Also note that $\tilde{\rho}_{\mathcal{S}_\infty}(x, y) < \infty$ for all $x, y \in \mathbb{Z}^d$ if and only if \mathcal{S}_∞ is connected.

The following lemma provides a preliminary bound on the shape of balls in \mathbb{Z}^d with respect to $\tilde{\rho}_{\mathcal{S}_\infty}$. For $x \in \mathbb{Z}^d$ and $r \geq 0$, let $B_{\mathcal{S}_\infty}(x, r)$ be the ball in \mathbb{Z}^d with center at x and radius $\lfloor r \rfloor$ with respect to $\tilde{\rho}_{\mathcal{S}_\infty}$, i.e.,

$$\tilde{B}_{\mathcal{S}_\infty}(x, r) = \left\{ y \in \mathbb{Z}^d : \tilde{\rho}_{\mathcal{S}_\infty}(x, y) \leq r \right\}. \quad (8.6)$$

Lemma 8.1. *Assume that the family of probability measures \mathbb{P}^u , $u \in (a, b)$, satisfies **P1** – **P3** and **S1** – **S2**. For any $u \in (a, b)$, there exists $C_\rho = C_\rho(u) < \infty$ such that for all $x \in \mathbb{Z}^d$ and $R \geq 1$,*

$$\mathbb{P}^u \left[B(x, R) \subseteq \tilde{B}_{\mathcal{S}_\infty}(x, C_\rho R) \right] \geq 1 - C_\rho e^{-(\log R)^{1+\Delta_{\mathcal{S}}}/C_\rho}, \quad (8.7)$$

where $\Delta_{\mathcal{S}} = \Delta_{\mathcal{S}}(u) > 0$ is defined in (1.6).

Proof of Lemma 8.1. By **P1**, (8.3), and (8.4), it suffices to prove the lemma for $x = 0 \in \mathbb{Z}^d$. Fix $u \in (a, b)$. By (1.8), there exist $c = c(u) > 0$ and $C = C(u) < \infty$ such that for all $y \in \mathbb{Z}^d$ and $R \geq 1$,

$$\mathbb{P}^u \left[|\Phi(y, \mathcal{S}_\infty) - y|_\infty > R \right] \leq C e^{-cfs(u, R)}. \quad (8.8)$$

Together with Theorem 1.3 and (1.6), (8.8) implies that there exists $C_\rho = C_\rho(u) < \infty$ such that for all $R \geq 1$,

$$\mathbb{P}^u \left[\exists z \in B(0, R) : \tilde{\rho}_{\mathcal{S}_\infty}(0, z) > C_\rho R \right] \leq C_\rho e^{-(\log R)^{1+\Delta_{\mathcal{S}}}/C_\rho}.$$

This implies (8.7) for $x = 0$, hence the proof of Lemma 8.1 is complete. \square

Now we state a variant of Theorem 1.5 in which $\rho_{\mathcal{S}}$ and $\mathbb{P}^u[\cdot | 0 \in \mathcal{S}_\infty]$ are replaced by $\tilde{\rho}_{\mathcal{S}_\infty}$ and \mathbb{P}^u , respectively.

Proposition 8.2. *Assume that the family of probability measures \mathbb{P}^u , $u \in (a, b)$, satisfies **P1** – **P3** and **S1** – **S2**. For any $u \in (a, b)$, there exists a convex compact subset $D_u \subset \mathbb{R}^d$ such that for all $0 < \varepsilon < 1$ there exists a \mathbb{P}^u -a.s. finite random variable $\tilde{R}_{\varepsilon, u}$ satisfying*

$$\forall R \geq \tilde{R}_{\varepsilon, u} : \mathbb{Z}^d \cap (1 - \varepsilon)R \cdot D_u \subseteq \tilde{B}_{\mathcal{S}_\infty}(0, R), \quad (8.9)$$

$$\forall R \geq \tilde{R}_{\varepsilon, u} : \mathbb{Z}^d \cap (1 + \varepsilon)R \cdot D_u \supseteq \tilde{B}_{\mathcal{S}_\infty}(0, R). \quad (8.10)$$

Before we prove Proposition 8.2 in Section 8.2, we deduce Theorem 1.5 from it.

Proof of Theorem 1.5. Let $u \in (a, b)$. We first observe that

$$\begin{aligned} \mathbb{P}^u [0 \in \mathcal{S}_\infty] &\stackrel{(8.9)}{=} \mathbb{P}^u \left[0 \in \mathcal{S}_\infty, \forall R \geq \tilde{R}_{\varepsilon, u} : \begin{array}{l} \mathcal{S}_\infty \cap (1 - \varepsilon)R \cdot D_u \subseteq \mathcal{S}_\infty \cap \tilde{B}_{\mathcal{S}_\infty}(0, R) \\ \mathcal{S}_\infty \cap (1 + \varepsilon)R \cdot D_u \supseteq \mathcal{S}_\infty \cap \tilde{B}_{\mathcal{S}_\infty}(0, R) \end{array} \right] \\ &\stackrel{(8.10)}{=} \mathbb{P}^u \left[0 \in \mathcal{S}_\infty, \forall R \geq \tilde{R}_{\varepsilon, u} : \begin{array}{l} \mathcal{S}_\infty \cap (1 - \varepsilon)R \cdot D_u \subseteq B_{\mathcal{S}}(0, R) \\ \mathcal{S}_\infty \cap (1 + \varepsilon)R \cdot D_u \supseteq B_{\mathcal{S}}(0, R) \end{array} \right]. \end{aligned}$$

Now the statement of Theorem 1.5 follows (with the same D_u and $\tilde{R}_{\varepsilon, u}$ as in Proposition 8.2) if we divide the above equality by $\mathbb{P}^u [0 \in \mathcal{S}_\infty]$ (which is positive by **S2**). \square

8.2 Proof of Proposition 8.2

Lemma 8.3. *Assume that the family of probability measures \mathbb{P}^u , $u \in (a, b)$, satisfies **P1** – **P3** and **S1** – **S2**. For any $u \in (a, b)$, there exists a unique norm $p_u(\cdot)$ on \mathbb{R}^d such that for any $x \in \mathbb{Z}^d$,*

$$p_u(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \tilde{\rho}_{S_\infty}(0, nx) \quad \mathbb{P}^u\text{-a.s.} \quad (8.11)$$

Proof. The proof is similar to the analogous result about the chemical distance on the unique infinite cluster of supercritical Bernoulli percolation in [13, Section 5, Remark (g)]. We use [14, Theorem 1] to deduce (8.11). In accordance with the notation of [14], we fix $u \in (a, b)$ and $x \in \mathbb{Z}^d$ and define

$$x_{m,n} := \tilde{\rho}_{S_\infty}(mx, nx), \quad m \leq n \in \mathbb{N}.$$

Now we check that the family $(x_{m,n} : m \leq n \in \mathbb{N})$ of random variables is a *subadditive process* for \mathbb{P}^u , i.e., it satisfies [14, Conditions **S**₁, **S**₂, **S**₃]. In our setting, these conditions correspond to (8.12), (8.13), and (8.14) below.

For $l \leq m \leq n \in \mathbb{N}$ the subadditivity condition

$$x_{l,n} \leq x_{l,m} + x_{m,n} \quad (8.12)$$

follows from the triangle inequality for $\tilde{\rho}_{S_\infty}(\cdot, \cdot)$.

By **P1** and (8.3), the joint law of $\tilde{\rho}_{S_\infty}$ -distances of the vertices of \mathbb{Z}^d under \mathbb{P}^u is translation invariant, in particular,

$$\text{the joint distribution of } (x_{m+1,n+1} : m \leq n \in \mathbb{N}) \text{ is the same as that of } (x_{m,n} : m \leq n \in \mathbb{N}), \quad (8.13)$$

moreover, the above shift is ergodic by **P1**.

It follows from (8.7) that for all $n \in \mathbb{N}$ we have

$$0 \leq \mathbb{E}^u[x_{0,n}] < \infty. \quad (8.14)$$

Having checked that $(x_{m,n} : m \leq n \in \mathbb{N})$ is a subadditive process for \mathbb{P}^u , we can use [14, Theorem 1] to obtain (8.11) with

$$p_u(x) = \inf_{n \in \mathbb{N}} \frac{1}{n} \mathbb{E}^u[\tilde{\rho}_{S_\infty}(0, nx)] = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}^u[\tilde{\rho}_{S_\infty}(0, nx)]. \quad (8.15)$$

Now we extend $p_u(\cdot)$ from \mathbb{Z}^d to \mathbb{R}^d , so that $p_u(\cdot)$ becomes a norm on \mathbb{R}^d . We first observe that the identity $p_u(nx) = np_u(x)$ follows from (8.11) for all $x \in \mathbb{Z}^d$ and $n \in \mathbb{N}$. This allows us to consistently extend $p_u(\cdot)$ to \mathbb{Q}^d by letting $p_u(x) = \frac{1}{n} p_u(nx)$ for any $x \in \mathbb{Q}^d$, where $n \in \mathbb{N}$ is chosen so that $nx \in \mathbb{Z}^d$.

By **P1**, the triangle inequality for $\tilde{\rho}_{S_\infty}(\cdot, \cdot)$, and (8.15), we obtain

$$p_u(-x) = p_u(x), \quad p_u(x + y) \leq p_u(x) + p_u(y), \quad x, y \in \mathbb{Q}^d. \quad (8.16)$$

By (8.8), (8.5), (8.14), (8.15), and (8.16), we have

$$|x|_1 \leq p_u(x) \leq \max_{1 \leq i \leq d} \mathbb{E}[\tilde{\rho}_{S_\infty}(0, e_i)] \cdot |x|_1, \quad x \in \mathbb{Q}^d. \quad (8.17)$$

Thus $p_u(\cdot)$ is a norm on \mathbb{Q}^d , which admits a unique continuous extension to \mathbb{R}^d . \square

In the sequel we denote by $Q(x, r)$ the closed L_∞ -ball in \mathbb{R}^d with radius $r \in \mathbb{R}_+$ and center $x \in \mathbb{R}^d$. Note that with the above definitions we have $B(x, r) = \mathbb{Z}^d \cap Q(x, r)$ for any $x \in \mathbb{Z}^d$ and $r \in \mathbb{R}_+$.

Proof of Proposition 8.2. Let $u \in (a, b)$. We define the compact convex set

$$D_u := \{x \in \mathbb{R}^d : p_u(x) \leq 1\} \quad (8.18)$$

using the norm $p_u(\cdot)$ from Lemma 8.3. We will show that (8.9) and (8.10) hold with this choice of D_u .

We first show the inclusion (8.9). Recall the constant C_ρ appearing in (8.7). For every $0 < \varepsilon < 1$ we choose $k = k(\varepsilon) \in \mathbb{N}$ and $x_1, \dots, x_k \in \mathbb{Q}^d \cap (1 - \varepsilon)D_u$ such that

$$(1 - \varepsilon)D_u \subseteq \bigcup_{j=1}^k \mathcal{Q}\left(x_j, \frac{\varepsilon}{3C_\rho}\right). \quad (8.19)$$

We fix $n \in \mathbb{N}$ such that $nx_j \in \mathbb{Z}^d$ for all $1 \leq j \leq k$ and define

$$x_j^R = \left\lfloor \frac{R}{n} \right\rfloor \cdot nx_j \in \mathbb{Z}^d, \quad 1 \leq j \leq k, \quad R \in \mathbb{N}. \quad (8.20)$$

Note that we have $\frac{1}{R}x_j^R \rightarrow x_j$ as $R \rightarrow \infty$ for every $1 \leq j \leq k$, thus there is a constant $C(\varepsilon)$ such that

$$\forall R \geq C(\varepsilon), 1 \leq j \leq k : \mathbb{Z}^d \cap \mathcal{Q}\left(Rx_j, \frac{\varepsilon}{3C_\rho}R\right) \subseteq \mathcal{B}\left(x_j^R, \frac{\varepsilon}{2C_\rho}R\right). \quad (8.21)$$

From Lemma 8.3 we obtain

$$\forall 1 \leq j \leq k : \lim_{R \rightarrow \infty} \frac{1}{R} \tilde{\rho}_{S_\infty}(0, x_j^R) \stackrel{\mathbb{P}^u\text{-a.s.}}{=} \frac{1}{n} p_u(nx_j) = p_u(x_j) \stackrel{(8.18)}{\leq} 1 - \varepsilon.$$

Thus, there exists a \mathbb{P}^u -almost surely finite random variable $\tilde{R}_{\varepsilon, u}^1$ such that (recall (8.6))

$$\forall R \geq \tilde{R}_{\varepsilon, u}^1, 1 \leq j \leq k : x_j^R \in \tilde{\mathcal{B}}_{S_\infty}\left(0, (1 - \frac{\varepsilon}{2})R\right). \quad (8.22)$$

By (8.7) and the Borel-Cantelli lemma, there exists a \mathbb{P}^u -almost surely finite random variable $\tilde{R}_{\varepsilon, u}^2$ such that

$$\forall R \geq \tilde{R}_{\varepsilon, u}^2, 1 \leq j \leq k : \mathcal{B}\left(x_j^R, \frac{\varepsilon}{2C_\rho}R\right) \subseteq \tilde{\mathcal{B}}_{S_\infty}\left(x_j^R, \frac{\varepsilon}{2}R\right). \quad (8.23)$$

Now we are ready to conclude the proof of (8.9). We have that \mathbb{P}^u -almost surely,

$$\begin{aligned} \forall R \geq \max\{\tilde{R}_{\varepsilon, u}^1, \tilde{R}_{\varepsilon, u}^2, C(\varepsilon)\} : \mathbb{Z}^d \cap (1 - \varepsilon)R \cdot D_u &\stackrel{(8.19)}{\subseteq} \mathbb{Z}^d \cap \bigcup_{j=1}^k \mathcal{Q}\left(Rx_j, \frac{\varepsilon}{3C_\rho}R\right) \\ &\stackrel{(8.21)}{\subseteq} \bigcup_{j=1}^k \mathcal{B}\left(x_j^R, \frac{\varepsilon}{2C_\rho}R\right) \stackrel{(8.23)}{\subseteq} \bigcup_{j=1}^k \tilde{\mathcal{B}}_{S_\infty}\left(x_j^R, \frac{\varepsilon}{2}R\right) \stackrel{(8.22)}{\subseteq} \tilde{\mathcal{B}}_{S_\infty}(0, R). \end{aligned} \quad (8.24)$$

It remains to show the inclusion (8.10). We first note that by (8.17) and (8.18),

$$D_u \subseteq \mathcal{Q}(0, 1).$$

For every $0 < \varepsilon < 1$, there exists $k \in \mathbb{N}$ and $x_1, \dots, x_k \in \mathbb{Q}^d \cap \mathcal{Q}(0, 2) \setminus (1 + \varepsilon)D_u$ such that

$$\mathcal{Q}(0, 2) \setminus (1 + \varepsilon)D_u \subseteq \bigcup_{j=1}^k \mathcal{Q}\left(x_j, \frac{\varepsilon}{3C_\rho}\right). \quad (8.25)$$

Let us fix $n \in \mathbb{N}$ such that $nx_j \in \mathbb{Z}^d$ for all $1 \leq j \leq k$. Define x_j^R by the formula (8.20). Similarly to (8.21), (8.22) and (8.23) there exists a constant $C'(\varepsilon) < \infty$ and \mathbb{P} -almost surely finite random variables $\tilde{R}_{\varepsilon,u}^3, \tilde{R}_{\varepsilon,u}^4$ such that

$$\forall R \geq C'(\varepsilon), 1 \leq j \leq k : \mathbb{Z}^d \cap \mathbb{Q} \left(Rx_j, \frac{\varepsilon}{3C_\rho} R \right) \subseteq \mathbb{B} \left(x_j^R, \frac{\varepsilon}{2C_\rho} R \right), \quad (8.26)$$

$$\forall R \geq \tilde{R}_{\varepsilon,u}^3, 1 \leq j \leq k : x_j^R \in \mathbb{Z}^d \setminus \tilde{\mathbb{B}}_{s_\infty} \left(0, (1 + \frac{\varepsilon}{2}) R \right), \quad (8.27)$$

$$\forall R \geq \tilde{R}_{\varepsilon,u}^4, 1 \leq j \leq k : \mathbb{B} \left(x_j^R, \frac{\varepsilon}{2C_\rho} R \right) \subseteq \tilde{\mathbb{B}}_{s_\infty} \left(x_j^R, \frac{\varepsilon}{2} R \right). \quad (8.28)$$

Similarly to (8.24), we obtain that \mathbb{P}^u -almost surely,

$$\begin{aligned} \forall R \geq \max\{\tilde{R}_{\varepsilon,u}^3, \tilde{R}_{\varepsilon,u}^4, C'(\varepsilon)\} : \mathbb{B}(0, 2R) \setminus (1 + \varepsilon)R \cdot D_u &\stackrel{(8.25)}{\subseteq} \mathbb{Z}^d \cap \bigcup_{j=1}^k \mathbb{Q} \left(Rx_j, \frac{\varepsilon}{3C_\rho} R \right) \stackrel{(8.26)}{\subseteq} \\ &\bigcup_{j=1}^k \mathbb{B} \left(x_j^R, \frac{\varepsilon}{2C_\rho} R \right) \stackrel{(8.28)}{\subseteq} \bigcup_{j=1}^k \tilde{\mathbb{B}}_{s_\infty} \left(x_j^R, \frac{\varepsilon}{2} R \right) \stackrel{(8.27)}{\subseteq} \mathbb{Z}^d \setminus \tilde{\mathbb{B}}_{s_\infty}(0, R). \end{aligned} \quad (8.29)$$

By (8.8), (8.5), and the Borel-Cantelli lemma, we obtain that there exists a \mathbb{P}^u -a.s. finite random variable $\tilde{R}_{\varepsilon,u}^5$ such that

$$\forall R \geq \tilde{R}_{\varepsilon,u}^5 : \tilde{\mathbb{B}}_{s_\infty}(0, R) \subseteq \mathbb{B}(0, 2R).$$

If we combine this observation with (8.29), we obtain (8.10), which concludes the proof of Proposition 8.2 and Theorem 1.5. \square

Acknowledgements. A large part of the work on this paper was done while the authors were at ETH Zürich. During that time, the research of AD was supported by an ETH Fellowship, and the research of BR and AS was supported by the grant ERC-2009-AdG 245728-RWPERCRI.

References

- [1] O. S. M. Alves, F. P. Machado and S. Yu Popov (2002) The Shape Theorem for the Frog Model. *Ann. Appl. Probab.* **12** (2), 533–546.
- [2] P. Antal and A. Pisztora (1996) On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24** (2), 1036–1048.
- [3] M. T. Barlow (2004) Random walks on supercritical percolation clusters. *Ann. Probab.* **32**, 3024–3084.
- [4] I. Benjamini and A.-S. Sznitman (2008) Giant component and vacant set for random walk on a discrete torus. *J. Eur. Math. Soc.* **10** (1), 133–172.
- [5] N. Berger and M. Biskup (2007) Quenched invariance principle for simple random walk on percolation cluster. *Probab. Theory Rel. Fields* **137**, 83–120.
- [6] J. Bricmont, J. L. Lebowitz and C. Maes (1987) Percolation in strongly correlated systems: the massless Gaussian field. *J. Stat. Phys.* **48** (5/6), 1249–1268.

- [7] S. R. Broadbent and J.M. Hammersley (1957) Percolation processes I. Crystals and mazes. *Proc. Camb. Phil. Soc.* **53**, 629–641.
- [8] J. Černý and S. Popov (2012) On the internal distance in the interlacement set. *Electron. J. Probab.* **17**, no. 29, 1–25.
- [9] A. Drewitz, B. Ráth and A. Sapozhnikov (2012) Local percolative properties of the vacant set of random interlacements with small intensity. arXiv:1206.6635.
- [10] N. Dunford and J.T. Schwartz (1958) *Linear operators*, Volume 1, Wiley-Interscience, New York.
- [11] O. Garet and R. Marchand (2004) Asymptotic shape for the chemical distance and first-passage percolation on the infinite Bernoulli cluster. *ESAIM: Probab. Stat.* **8**, 169–199.
- [12] G.R. Grimmett (1999) *Percolation*. Springer-Verlag, Berlin, second edition.
- [13] G.R. Grimmett and J.M. Marstrand (1990) The supercritical phase of percolation is well behaved. *Proc. Roy. Soc. London Ser. A* **430**, 439–457.
- [14] J. F. C. Kingman (1973) Subadditive ergodic theory. *Ann. Probab.* **1**, 883–909.
- [15] H. Lacoïn and J. Tykesson (2012) On the easiest way to connect k points in the random interlacements process. arXiv:1206.4216.
- [16] P. Mathieu and E. Remy (2004) Isoperimetry and heat kernel decay on percolations clusters. *Ann. Probab.* **32**, 100–128.
- [17] P. Mathieu and A. L. Piatnitski (2007) Quenched invariance principles for random walks on percolation clusters. *Proceedings of the Royal Society A* **463**, 2287–2307.
- [18] E. Procaccia and E. Shellef (2012) On the range of a random walk in a torus and random interlacements. arXiv:1007.1401v2.
- [19] E.B. Procaccia and J. Tykesson (2011) Geometry of the random interlacement. *Electron. Commun. Probab.* **16**, 528–544.
- [20] B. Ráth and A. Sapozhnikov (2012) Connectivity properties of random interlacement and intersection of random walks. *ALEA* **9**, 67–83.
- [21] B. Ráth and A. Sapozhnikov (2011) On the transience of random interlacements. *Electron. Commun. Probab.* **16**, 379–391.
- [22] B. Ráth and A. Sapozhnikov (2011) The effect of small quenched noise on connectivity properties of random interlacements. arXiv:1109.5086.
- [23] P. F. Rodriguez and A.-S. Sznitman (2012) Phase transition and level-set percolation for the Gaussian free field. To appear in *Comm. Math. Phys.* arXiv:1202.5172.
- [24] V. Sidoravicius and A.-S. Sznitman (2004) Quenched invariance principles for walks on clusters of percolation or among random conductances. *Prob. Theory Rel. Fields* **129**, 219–244.
- [25] V. Sidoravicius and A.-S. Sznitman (2009) Percolation for the Vacant Set of Random Interlacements. *Comm. Pure Appl. Math.* **62** (6), 831–858.

- [26] A.-S. Sznitman (2010) Vacant set of random interlacements and percolation. *Ann. Math.* **171** (2), 2039–2087.
- [27] A.-S. Sznitman (2012) Decoupling inequalities and interlacement percolation on $G \times \mathbb{Z}$. *Invent. Math.* **187** (3), 645–706.
- [28] A. Teixeira (2009) On the uniqueness of the infinite cluster of the vacant set of random interlacements. *Ann. Appl. Probab.* **19** (1), 454–466.
- [29] A. Teixeira (2011) On the size of a finite vacant cluster of random interlacements with small intensity. *Probab. Theory Related Fields* **150** (3–4), 529–574.
- [30] A. Teixeira and D. Windisch (2011) On the fragmentation of a torus by random walk. *Comm. Pure Appl. Math.* **64** (12), 1599–1646.
- [31] D. Windisch (2008) Random walk on a discrete torus and random interlacements. *Electron. Commun. Probab.* **13**, 140–150.